

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
25. März 2004 (25.03.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/024933 A2

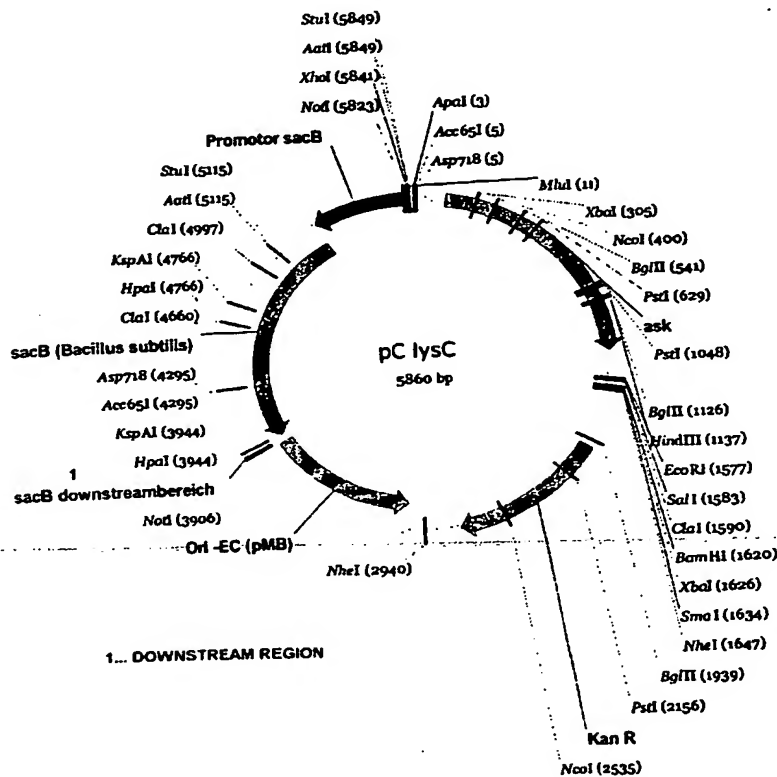
- (51) Internationale Patentklassifikation⁷: C12P 13/04, 13/12
- (21) Internationales Aktenzeichen: PCT/EP2003/009453
- (22) Internationales Anmeldedatum:
26. August 2003 (26.08.2003)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:
102 39 082.7 26. August 2002 (26.08.2002) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; 67056 Ludwigshafen (DE).

- (72) Erfinder; und
(75) Erfinder/Anmelder (nur für US): KRÖGER, Burkhard [DE/DE]; Im Waldhof 1, 67117 Limburgerhof (DE). ZELDER, Oskar [DE/DE]; Franz-Stützel-Str. 8, 67346 Speyer (DE). KLOPPROGGE, Corinna [DE/DE]; Rastatter Str. 10, 68239 Mannheim (DE). SCHRÖDER, Hartwig [DE/DE]; Benzstr. 4, 69226 Nussloch (DE). HÄFNER, Stefan [DE/DE]; Luitpoldstr. 11, 67063 Ludwigshafen (DE).
- (74) Anwalt: KINZEBACH, Werner; Reitstötter, Kinzebach & Partner (GbR), Sternwartstr. 4, 81679 München (DE).
- (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR ZYMOTIC PRODUCTION OF FINE CHEMICALS (METY) CONTAINING SULPHUR

(54) Bezeichnung: VERFAHREN ZUR FERMENTATIVEN HERSTELLUNG SCHWEFELHALTIGER FEINCHEMIKALIEN (METY)



(57) Abstract: The invention relates to methods for the zymotic production of fine chemicals, especially L-methionine, containing sulphur using bacteria, wherein a nucleotide sequence coding for a methionine-synthase (methH)-gene is expressed.

(57) Zusammenfassung: Die Erfindung betrifft Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für ein Methionin-Synthase (methH)-Gen kodierende Nukleotidsequenz exprimiert wird.



KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Rechenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

(84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL,

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

VERFAHREN ZUR FERMENTATIVEN HERSTELLUNG SCHWEFELHALTIGER FEINCHEMIKALIEN (METY)

Beschreibung

5 Gegenstand der Erfindung ist ein Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für ein O-Acetyl-Homoserin-Sulphydrolase (metY)-Gen kodierende Nukleotidsequenz exprimiert wird.

Stand der Technik

10 Schwefelhaltige Feinchemikalien, wie zum Beispiel Methionin, Homocystein, S-Adenosyl-Methionin, Glutathion, Cystein, Biotin, Thiamin, Liponsäure werden über natürliche Stoffwechselprozesse in Zellen hergestellt und werden in vielen Industriezweigen verwendet, einschließ-
15 lich der Nahrungsmittel-, Futtermittel-, Kosmetik- und pharmazeutischen Industrie. Diese Substanzen, die zusammen als "schwefelhaltige Feinchemikalien" bezeichnet werden, umfassen organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Vitamine
20 und Cofaktoren. Ihre Produktion erfolgt am zweckmäßigsten im Großmaßstab mittels Anzucht von Bakterien, die entwickelt wurden, um große Mengen der jeweils gewünschten Substanz zu produzieren und sezernieren. Für diesen Zweck besonders geeignete Organismen sind coryne-
25 forme Bakterien, gram-positive nicht-pathogene Bakterien.

30 Es ist bekannt, dass Aminosäuren durch Fermentation von Stämmen coryneformer Bakterien, insbesondere Corynebacterium glutamicum, hergestellt werden. Wegen der großen Bedeutung wird ständig an der Verbesserung der Herstellverfahren gearbeitet. Verfahrensverbesserungen können fermentationstechnische Maßnahmen, wie zum Beispiel Rührung und Versorgung mit
35 Sauerstoff, oder die Zusammensetzung der Nährmedien, wie zum Beispiel die Zuckerkonzentration während der Fermentation, oder die Aufarbeitung zum Produkt, beispielsweise durch Ionenaustauschchromatographie, oder die intrinsischen Leistungseigenschaften des Mikroorganismus selbst betreffen.

40 Über Stammselektion sind eine Reihe von Mutantenstämmen entwickelt worden, die ein Sortiment wünschenswerter Verbindungen aus der Reihe der schwefelhaltigen Feinchemikalien produzieren. Zur Verbesserung der Leistungseigenschaften dieser Mikroorganismen hinsichtlich der Produktion eines bestimmten Moleküls werden Methoden der Mutagenese, Selektion und Mutantenauswahl angewendet. Dies ist jedoch ein zeitaufwendiges und schwieriges Verfahren. Auf
45 diese Weise erhält man z.B. Stämme, die resistent gegen Antimetabolite, wie z. B. die Methionin-Analoga α -Methyl-Methionin, Ethionin, Norleucin, N-Acetylnorleucin, S-Trifluoromethyl-

homocystein, 2-Amino-5-heptenoinsäure, Seleno-Methionin, Methioninsulfoximin, Methoxin, 1-Aminocyclopentan-Carboxylsäure oder auxotroph für regulatorisch bedeutsame Metabolite sind und schwefelhaltige Feinchemikalien, wie z. B. L-Methionin, produzieren.

- 5 Seit einigen Jahren werden ebenfalls Methoden der rekombinanten DNA-Technik zur Stammverbesserung von L-Aminosäure produzierender Stämme von *Corynebacterium* eingesetzt, indem man einzelne Aminosäure-Biosynthesegene amplifiziert und die Auswirkung auf die Aminosäure-Produktion untersucht.
- 10 Die WO-A-02/18613 beschreibt die Nukleinsäure- und Aminosäuresequenz für metY aus *C. glutamicum* und dessen Verwendung zur Herstellung von L-Lysin.

Kurze Beschreibung der Erfindung

- 15 Der Erfindung lag die Aufgabe zugrunde, ein neues Verfahren zur verbesserten fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, bereitzustellen.

- Gelöst wird obige Aufgabe durch Bereitstellung eines Verfahrens zur fermentativen Herstellung einer schwefelhaltigen Feinchemikalie, umfassend die Expression einer heterologen Nukleotidsequenz, welche für ein Protein mit metY-Aktivität kodiert, in einem coryneformen Bakterium.
- 20

Ein erster Gegenstand der Erfindung ist Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen Feinchemikalie, welches folgende Schritte umfasst:

- 25 a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie produzierenden coryneformen Bakterienkultur, wobei in den coryneformen Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird, welche für ein Protein mit O-Acetyl-Homoserin-Sulphydrolase (metY) –Aktivität kodiert;
- b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den Zellen der Bakterien, und
- 30 c) Isolieren der schwefelhaltigen Feinchemikalie, welche vorzugsweise L-Methionin umfasst.

- Vorzugsweise besitzt obige heterologe metY-kodierende Nukleotidsequenz zur metY-kodierenden Sequenz aus *Corynebacterium glutamicum* ATCC 13032 eine Sequenzhomologie von weniger als 100%, wie z.B. mehr als 70%, wie 75, 80, 85, 90 oder 95 %, oder weniger als 70%, wie z.B. bis zu 60, 50, 40, 30, 20 oder 10 %. Die metY-kodierende Sequenz ist vorzugsweise aus einem der folgenden Organismen von Liste I abgeleitet:
- 35

Liste I

<i>Corynebacterium diptheriae</i>	ATCC 14779
<i>Mycobacterium tuberculosis</i> CDC1551	ATCC 25584
<i>Clostridium acetobutylicum</i>	ATCC 824
<i>Bacillus halodurans</i>	ATCC21591
<i>Bacillus stearothermophilus</i>	ATCC 12980
<i>Chlorobium tepidum</i>	ATCC 49652
<i>Synechococcus</i> sp.	ATCC27104
<i>Emericella nidulans</i>	ATCC 36104
<i>Bacteroides fragilis</i>	ATCC 25285
<i>Lactococcus lactis</i>	ATCC 7962
<i>Bordetella bronchiseptica</i>	ATCC 19395
<i>Pseudomonas aeruginosa</i>	ATCC 17933
<i>Nitrosomonas europaea</i>	ATCC 19718
<i>Sinorhizobium meliloti</i>	ATCC 4399
<i>Thermotoga maritima</i>	ATCC 43589
<i>Streptococcus mutans</i>	ATCC 25175
<i>Burkholderia cepacia</i>	ATCC 25416
<i>Deinococcus radiodurans</i>	ATCC 13939
<i>Rhodobacter capsulatus</i>	ATCC 11166
<i>Pasteurella multocida</i>	ATCC 6530
<i>Clostridium difficile</i>	ATCC 9689
<i>Campylobacter jejuni</i>	ATCC 33560
<i>Streptococcus pneumoniae</i>	ATCC 6308
<i>Saccharomyces cerevisiae</i>	ATCC 2704
<i>Kluyveromyces lactis</i>	ATCC 8585
<i>Candida albicans</i>	ATCC 10231
<i>Schizosaccharomyces pombe</i>	ATCC 24969

5 ATCC: American Type Culture Collection, Rockville, MD, USA

Die erfindungsgemäß eingesetzte metY-kodierende Sequenz umfasst vorzugsweise eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 oder eine dazu homologe Nukleotidsequenz, welche für ein

10 Protein mit metY-Aktivität kodiert.

Die erfindungsgemäß eingesetzte metY-kodierende Sequenz kodiert außerdem vorzugsweise für ein Protein mit metY-Aktivität, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52

und 54 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metY-Aktivität steht, umfasst.

- Die kodierende metY-Sequenz ist vorzugsweise eine in coryneformen Bakterien replizierbare
5 oder eine stabil in das Chromosom integrierte DNA oder eine RNA.

Gemäß einer bevorzugten Ausführungsform wird das erfindungsgemäße Verfahren durchgeführt, indem man

- 10 a) einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden metY-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
b) einen Stamm einsetzt, in dem die kodierende metY-Sequenz in das Chromosom des Bakteriums integriert wurde

- 15 Es ist weiterhin bevorzugt, die kodierende metY-Sequenz für die Fermentation zu überexprimieren.

- Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie
20 verstärkt ist; und / oder
in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet sind, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.

- Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie
25 durch Stoffwechselmetabolite in seiner Aktivität nicht in unerwünschter Weise beeinflusst wird.

- Gemäß einer weiteren Ausführungsform des erfindungsgemäßen Verfahrens werden deshalb coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt
30 unter

- a) dem für eine Aspartatkinase kodierenden Gen *lysC*,
b) dem für eine Aspartat-Semialdehyd-Dehydrogenase kodierenden Gen *asd*
c) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen *gap*,
35 d) dem für die 3-Phosphoglycerat Kinase kodierenden Gen *pgk*,
e) dem für die Pyruvat Carboxylase kodierenden Gen *pyc*,
f) dem für die Triosephosphat Isomerase kodierenden Gen *tpi*,

- g) dem für die Homoserin O-Acetyltransferase kodierenden Gen metA,
h) dem für die Cystathionin-gamma-Synthase kodierenden Gen metB,
i) dem für die Cystathionin-gamma-Lyase kodierenden Gen metC,
j) dem für die Serin-Hydroxymethyltransferase kodierenden Gen glyA,
5 k) dem für die Methionin Synthase kodierenden Gen methI,
l) dem für die Methylen-Tetrahydrofolat-Reduktase kodierenden Gen, metF
m) dem für die Phosphoserin-Aminotransferase kodierenden Gen serC
n) dem für die Phosphoserin-Phosphatase kodierenden Gen serB,
o) dem für die Serine Acetyl-Transferase kodierenden Gen cysE,
10 p) dem für die Homoserin-Dehydrogenase kodierenden Gen hom,
überexprimiert ist.

15 Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene ausgewählt unter Genen der oben genannten Gruppe a) bis p) mutiert ist, so dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden und dass insbesondere die erfindungsgemäße Produktion der Feinchemikalie nicht beeinträchtigt wird.

20 Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- q) dem für die Homoserine-Kinase kodierenden Gen thrB,
r) dem für die Threonin Dehydratase kodierenden Gen ilvA,
s) dem für die Threonin Synthase kodierenden Gen thrC
25 t) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen ddh
u) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen pck,
v) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen pgi,
w) dem für die Pyruvat-Oxidase kodierenden Gen poxB,
x) dem für die Dihydrodipicolinat Synthase kodierenden Gen dapA,
30 y) dem für die Dihydrodipicolinat Reduktase kodierenden Gen dapB; oder
z) dem für die Diaminopicolinat Decarboxylase kodierenden Gen lysA

abschwächt ist, insbesondere durch Verringerung der Expressionsrate des korrespondierenden Gens.

35 Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene der obigen Gruppen

q) bis z) mutiert ist, so dass die enzymatische Aktivität des korrespondierenden Proteins teilweise oder vollständig verringert wird.

5 Vorzugsweise werden in dem erfindungsgemäßen Verfahren Mikroorganismen der Art *Corynebacterium glutamicum* eingesetzt.

Ein weiterer Gegenstand der Erfindung betrifft ein Verfahren zur Herstellung eines L-Methionin-haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst

- 10 a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
- b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
- c) Entfernung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
- 15 d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.

Gegenstand der Erfindung sind ebenfalls die erstmalig aus obigen Mikroorganismen isolierten kodierenden metY-Sequenzen, die davon kodierten O-Acetyl-Homoserin-Sulphydrolasen sowie die funktionalen Homologen dieser Polynukleotide bzw. Proteine.

20

Detaillierte Beschreibung der Erfindung

a) Allgemeine Begriffe

25 Als Proteine mit der Aktivität der O-Acetyl-Homoserin Sulphydrolase auch metY (EC 4.2.99.10) genannt, werden solche Proteine beschrieben, die in der Lage sind O-Acetyl-Homoserin und Sulfid unter Verwendung des Cofaktors Pyridoxal-Phosphat zu Homocystein umzusetzen. Der Fachmann unterscheidet die Aktivität der O-Acetyl-Homoserin Sulphydrolase von der O-Succinyl-Homoserin-Sulphydrolase auch metZ genannt. In dem letztgenannten Enzym dient O-

30 Succinyl-Homoserin und nicht O-Acetyl-Homoserin als Substrat der Reaktion. Der Fachmann kann die enzymatische Aktivität von metY durch Enzymtests nachweisen, Vorschriften dafür können sein: Shimizu H. Yamagata S. Masui R. Inoue Y. Shibata T. Yokoyama S. Kuramitsu S. Iwama T. Biochimica et Biophysica Acta. 1549(1):61-72, 2001, Yamagata S. Isaji M. Nakamura K. Fujisaki S. Doi K. Bawden S. D'Andrea R. Applied Microbiology & Biotechnology. 42(1):92-9,

35 1994

Im Rahmen der vorliegenden Erfindung umfasst der Begriff „schwefelhaltige Feinchemikalie“ jegliche chemische Verbindung, die wenigstens ein Schwefelatom kovalent gebunden enthält und durch ein erfindungsgemäßes Fermentationsverfahrens zugänglich ist. Nichtlimitierende Beispiele dafür sind Methionin, Homocystein, S-Adenosyl-Methionin, insbesondere Methionin, und S-Adenosyl-Methionin.

Im Rahmen der vorliegenden Erfindung umfassen die Begriffe „L-Methionin“, „Methionin“, Homocystein und S-Adenosylmethionin auch die korrespondierenden Salze, wie z. B. Methionin-Hydrochlorid oder Methionin-Sulfat.

„Polynukleotide“ bezeichnet im allgemeinen Polyribonukleotide (RNA) und Polydeoxyribonukleotide (DNA), wobei es sich um nicht modifizierte RNA oder DNA oder modifizierte RNA oder DNA handeln kann.

Unter „Polypeptiden“ versteht man erfindungsgemäß Peptide oder Proteine, die zwei oder mehr über Peptidbindungen verbundene Aminosäuren enthalten.

Der Begriff „Stoffwechselmetabolit“ bezeichnet chemische Verbindungen, die im Stoffwechsel von Organismen als Zwischen- oder auch Endprodukte vorkommen und die neben ihrer Eigenschaft als chemische Bausteine auch modulierende Wirkung auf Enzyme und ihre katalytische Aktivität haben können. Dabei ist aus der Literatur bekannt, dass solche Stoffwechselmetabolite sowohl hemmend als auch stimulierend auf die Aktivität von Enzymen wirken können (Biochemistry, Stryer, Lubert, 1995 W. H. Freeman & Company, New York, New York.). In der Literatur ist auch beschrieben, dass es möglich ist durch Maßnahmen wie Mutation der genomischen DNA durch UV-Strahlung, ionisierender Strahlung oder mutagene Substanzen und nachfolgender Selektion auf bestimmte Phänotypen in Organismen solche Enzyme zu produzieren, in denen die Beeinflussung durch Stoffwechselmetabolite verändert wurde (Sahm H. Eggeling L. de Graaf AA. Biological Chemistry 381(9-10):899-910, 2000; Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94). Diese veränderten Eigenschaften können auch durch gezielte Maßnahmen erreicht werden. Dabei ist dem Fachmann bekannt, dass in Genen für Enzyme bestimmte Nukleotide der für das Protein kodierenden DNA gezielt verändert werden können, so dass das aus der exprimierten DNA-Sequenz resultierende Protein bestimmte neue Eigenschaften aufweist, so zum Beispiel, dass die modulierende Wirkung von Stoffwechselmetaboliten gegenüber dem nicht veränderten Protein verändert ist

Enzyme können derart in ihrer Aktivität beeinflusst werden, dass es zu einer Verringerung der Reaktionsgeschwindigkeit, oder zu einer Veränderung der Affinität gegenüber dem Substrat o-

der zu einer Änderung der Reaktionsgeschwindigkeiten kommt.

Die Begriffe "exprimieren" bzw. "Verstärkung" oder „Überexpression“ beschreiben im Kontext der Erfindung die Produktion bzw. Erhöhung der intrazellulären Aktivität eines oder mehrerer Enzyme in einem Mikroorganismus, die durch die entsprechende DNA kodiert werden. Dazu kann man beispielsweise ein Gen in einen Organismus einbringen, ein vorhandenes Gen durch ein anderes Gen ersetzen, die Kopienzahl des Gens bzw. der Gene erhöhen, einen starken Promotor verwenden oder ein Gen verwenden, das für ein entsprechendes Enzym mit einer hohen Aktivität kodiert und man kann gegebenenfalls diese Maßnahmen kombinieren.

b) Erfindungsgemäße metY-Proteine

Erfindungsgemäß mit umfasst sind ebenfalls „funktionale Äquivalente“ der konkret offenbarten metY-Enzyme aus Organismen obiger Liste I.

„Funktionale Äquivalente“ oder Analoga der konkret offenbarten Polypeptide sind im Rahmen der vorliegenden Erfindung davon verschiedene Polypeptide, welche weiterhin die gewünschte biologische Aktivität, wie z.B. Substratspezifität, besitzen.

Unter "funktionalen Äquivalenten" versteht man erfindungsgemäß insbesondere Mutanten, welche in wenigstens einer der oben genannten Sequenzpositionen eine andere als die konkret genannte Aminosäure aufweisen aber trotzdem eine der oben genannten biologische Aktivitäten besitzen. "Funktionale Äquivalente" umfassen somit die durch eine oder mehrere Aminosäure-Additionen, -Substitutionen, -Deletionen und/oder -Inversionen erhältlichen Mutanten, wobei die genannten Veränderungen in jeglicher Sequenzposition auftreten können, solange sie zu einer Mutante mit dem erfindungsgemäßen Eigenschaftsprofil führen. Funktionale Äquivalenz ist insbesondere auch dann gegeben, wenn die Reaktivitätsmuster zwischen Mutante und unverändertem Polypeptid qualitativ übereinstimmen, d.h. beispielsweise gleiche Substrate mit unterschiedlicher Geschwindigkeit umgesetzt werden.

"Funktionale Äquivalente" umfassen natürlich auch Polypeptide welche aus anderen Organismen zugänglich sind, sowie natürlich vorkommende Varianten. Beispielsweise lassen sich durch Sequenzvergleich Bereiche homologer Sequenzregionen festlegen und in Anlehnung an die konkreten Vorgaben der Erfindung äquivalente Enzyme ermitteln.

„Funktionale Äquivalente“ umfassen ebenfalls Fragmente, vorzugsweise einzelne Domänen oder Sequenzmotive, der erfindungsgemäßen Polypeptide, welche z.B. die gewünschte biologi-

sche Funktion aufweisen.

5 „Funktionale Äquivalente“ sind außerdem Fusionsproteine, welche ein der oben genannten Polypeptidsequenzen oder davon abgeleitete funktionale Äquivalente und wenigstens eine weitere, davon funktionell verschiedene, heterologe Sequenz in funktioneller N- oder C-terminaler Verknüpfung (d.h. ohne gegenseitigen wesentliche funktionelle Beeinträchtigung der Fusionsproteinteile) aufweisen. Nichtlimitierende Beispiele für derartige heterologe Sequenzen sind z.B. Signalpeptide, Enzyme, Immunoglobuline, Oberflächenantigene, Rezeptoren oder Rezeptorliganden.

10 Erfindungsgemäß mit umfasste „funktionale Äquivalente“ sind Homologe zu den konkret offenbarten Proteinen. Diese besitzen wenigstens 30%, oder etwa 40%, 50 %, vorzugsweise wenigstens etwa 60 %, 65%, 70%, oder 75% ins besondere wenigsten 85 %, wie z.B. 90%, 95% oder 99%, Homologie zu einer der konkret offenbarten Sequenzen, berechnet nach dem Algorithmus von Pearson und Lipman, Proc. Natl. Acad. Sci. (USA) 85(8), 1988, 2444-2448.

15 Homologe der erfindungsgemäßen Proteine oder Polypeptide können durch Mutagenese erzeugt werden, z.B. durch Punktmutation oder Verkürzung des Proteins. Der Begriff „Homolog“, wie er hier verwendet wird, betrifft eine variante Form des Proteins, die als Agonist oder Antagonist der Protein-Aktivität wirkt.

20 Homologe des erfindungsgemäßen Proteine können durch Screening kombinatorischer Banken von Mutanten, wie z.B. Verkürzungsmutanten, identifiziert werden. Beispielsweise kann eine variierte Bank von Protein-Varianten durch kombinatorische Mutagenese auf Nukleinsäureebene erzeugt werden, wie z.B. durch enzymatisches Ligieren eines Gemisches synthetischer Oligonukleotide. Es gibt eine Vielzahl von Verfahren, die zur Herstellung von Banken potentieller Homologer aus einer degenerierten Oligonukleotidsequenz verwendet werden können. Die chemische Synthese einer degenerierten Gensequenz kann in einem DNA-Syntheseautomaten durchgeführt werden, und das synthetische Gen kann dann in einen geeigneten Expressionsvektor ligiert werden. Die Verwendung eines degenerierten Gensatzes ermöglicht die Bereitstellung sämtlicher Sequenzen in einem Gemisch, die den gewünschten Satz an potentiellen Proteinsequenzen codieren. Verfahren zur Synthese degenerierter Oligonukleotide sind dem Fachmann bekannt (z.B. Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al. (1983) Nucleic Acids Res. 11:477).

Zusätzlich können Banken von Fragmenten des Protein-Codons verwendet werden, um eine variierte Population von Protein-Fragmenten zum Screening und zur anschließenden Selektion von Homologen eines erfindungsgemäßen Proteins zu erzeugen. Bei einer Ausführungsform kann eine Bank von kodierenden Sequenzfragmenten durch Behandeln eines doppelsträngigen PCR-Fragmentes einer kodierenden Sequenz mit einer Nuklease unter Bedingungen, unter denen ein Nicking nur etwa einmal pro Molekül erfolgt, Denaturieren der doppelsträngigen DNA, Renaturieren der DNA unter Bildung doppelsträngiger DNA, die Sense-/Antisense-Paare von verschiedenen genickten Produkten umfassen kann, Entfernen einzelsträngiger Abschnitte aus neu gebildeten Duplices durch Behandlung mit S1-Nuclease und Ligieren der resultierenden Fragmentbank in einen Expressionsvektor erzeugt werden. Durch dieses Verfahren kann eine Expressionsbank hergeleitet werden, die N-terminale, C-terminale und interne Fragmente mit verschiedenen Größen des erfindungsgemäßen Proteins kodiert.

Im Stand der Technik sind mehrere Techniken zum Screening von Genprodukten kombinatorischer Banken, die durch Punktmutationen oder Verkürzung hergestellt worden sind, und zum Screening von cDNA-Banken auf Genprodukte mit einer ausgewählten Eigenschaft bekannt. Diese Techniken lassen sich an das schnelle Screening der Genbanken anpassen, die durch kombinatorische Mutagenese erfindungsgemäßer Homologer erzeugt worden sind. Die am häufigsten verwendeten Techniken zum Screening großer Genbanken, die einer Analyse mit hohem Durchsatz unterliegen, umfassen das Klonieren der Genbank in replizierbare Expressionsvektoren, Transformieren der geeigneten Zellen mit der resultierenden Vektorenbank und Exprimieren der kombinatorischen Gene unter Bedingungen, unter denen der Nachweis der gewünschten Aktivität die Isolation des Vektors, der das Gen codiert, dessen Produkt nachgewiesen wurde, erleichtert. Recursive-Ensemble-Mutagenese (REM), eine Technik, die die Häufigkeit funktionaler Mutanten in den Banken vergrößert, kann in Kombination mit den Screeningtests verwendet werden, um Homologe zu identifizieren (Arkin und Yourvan (1992) PNAS 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331

c) Erfindungsgemäße Polynukleotide

Gegenstand der Erfindung sind ebenso Nukleinsäuresequenzen (einzel- und doppelsträngige DNA- und RNA-Sequenzen, wie z.B. cDNA und mRNA), kodierend für eines der obigen metY-Enzyme und deren funktionalen Äquivalenten, welche z.B. auch unter Verwendung künstlicher Nukleotidanaloge zugänglich sind.

Die Erfindung betrifft sowohl isolierte Nukleinsäuremoleküle, welche für erfindungsgemäße Polypeptide bzw. Proteine oder biologisch aktive Abschnitte davon kodieren, sowie Nukleinsäure-

fragmente, die z.B. zur Verwendung als Hybridisierungs sonden oder Primer zur Identifizierung oder Amplifizierung von erfindungsgemäßer kodierenden Nukleinsäuren verwendet werden können.

- 5 Die erfindungsgemäßen Nukleinsäuremoleküle können zudem untranslatierte Sequenzen vom 3'- und/oder 5'-Ende des kodierenden Genbereichs enthalten

- 10 Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure zugegen sind und kann überdies im wesentlichen frei von anderem zellulären Material oder Kulturmedium sein, wenn es durch rekombinante Techniken hergestellt wird, oder frei von chemischen Vorstufen oder anderen Chemikalien sein, wenn es chemisch synthetisiert wird.

- 15 Die Erfindung umfasst weiterhin die zu den konkret beschriebenen Nukleotidsequenzen komplementären Nukleinsäuremoleküle oder einen Abschnitt davon.

- 20 Die erfindungsgemäß Nukleotidsequenzen ermöglichen die Erzeugung von Sonden und Primern, die zur Identifizierung und/oder Klonierung von homologer Sequenzen in anderen Zelltypen und Organismen verwendbar sind. Solche Sonden bzw. Primer umfassen gewöhnlich einen Nukleotidsequenzbereich, der unter stringenten Bedingungen an mindestens etwa 12, vorzugsweise mindestens etwa 25, wie z.B. etwa 40, 50 oder 75 aufeinanderfolgende Nukleotide eines Sense-Stranges einer erfindungsgemäßen Nukleinsäuresequenz oder eines entsprechenden Antisense-Stranges hybridisiert.

- 25 Weitere erfindungsgemäße Nukleinsäuresequenzen sind abgeleitet von SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 oder 53 und unterscheiden sich davon durch Addition, Substitution, Insertion oder Deletion einzelner oder mehrerer Nukleotide, kodieren aber weiterhin für Polypeptide mit dem gewünschten Eigenschaftsprofil. Dies können Polynukleotide sein, die zu obigen Sequenzen in mindestens etwa 50%, 55%, 30 60%, 65%, 70%, 80% oder 90%, vorzugsweise in mindestens etwa 95%, 96%, 97%, 98% oder 99% der Sequenzpositionen identisch sind.

- 35 Erfindungsgemäß umfasst sind auch solche Nukleinsäuresequenzen, die sogenannte stumme Mutationen umfassen oder entsprechend der Codon-Nutzung eines speziellen Ursprungs- oder Wirtsorganismus, im Vergleich zu einer konkret genannten Sequenz verändert sind, ebenso wie natürlich vorkommende Varianten, wie z.B. Spleißvarianten oder Allelvarianten, davon. Gegenstand sind ebenso durch konservative Nukleotids substitutionen (d.h. die betreffende Aminosäure

wird durch eine Aminosäure gleicher Ladung, Größe, Polarität und/oder Löslichkeit ersetzt) erhältliche Sequenzen.

- 5 Gegenstand der Erfindung sind auch die durch Sequenzpolymorphismen von den konkret offenbarten Nukleinsäuren abgeleiteten Moleküle. Diese genetischen Polymorphismen können zwischen Individuen innerhalb einer Population aufgrund der natürlichen Variation existieren. Diese natürlichen Variationen bewirken üblicherweise eine Varianz von 1 bis 5 % in der Nukleotidsequenz eines Gens.
- 10 Weiterhin umfasst die Erfindung auch Nukleinsäuresequenzen, welchen mit oben genannten kodierenden Sequenzen hybridisieren oder dazu komplementär sind. Diese Polynukleotide lassen sich bei Durchmusterung von genomischen oder cDNA-Banken auffinden und gegebenenfalls daraus mit geeigneten Primern mittels PCR vermehren und anschließend beispielsweise mit geeigneten Sonden isolieren. Eine weitere Möglichkeit bietet die Transformation geeigneter
- 15 Mikroorganismen mit erfindungsgemäßen Polynukleotiden oder Vektoren, die Vermehrung der Mikroorganismen und damit der Polynukleotide und deren anschließende Isolierung. Darüber hinaus können erfindungsgemäße Polynukleotide auch auf chemischem Wege synthetisiert werden.
- 20 Unter der Eigenschaft, an Polynukleotide „hybridisieren“ zu können, versteht man die Fähigkeit eines Poly- oder Oligonukleotids unter stringenten Bedingungen an eine nahezu komplementäre Sequenz zu binden, während unter diesen Bedingungen unspezifische Bindungen zwischen nicht-komplementären Partnern unterbleiben. Dazu sollten die Sequenzen zu 70-100%, vorzugsweise zu 90-100%, komplementär sein. Die Eigenschaft komplementärer Sequenzen, spezifisch aneinander binden zu können, macht man sich beispielsweise in der Northern- oder Southern-Blot-Technik oder bei der Primerbindung in PCR oder RT-PCR zunutze. Üblicherweise werden dazu Oligonukleotide ab einer Länge von 30 Basenpaaren eingesetzt. Unter stringenten Bedingungen versteht man beispielsweise in der Northern-Blot-Technik die Verwendung einer
- 25 50 – 70 °C, vorzugsweise 60 – 65 °C warmen Waschlösung, beispielsweise 0,1x SSC-Puffer mit
- 30 0,1% SDS (20x SSC: 3M NaCl, 0,3M Na-Citrat, pH 7,0) zur Elution unspezifisch hybridisierter cDNA-Sonden oder Oligonukleotide. Dabei bleiben, wie oben erwähnt, nur in hohem Maße komplementäre Nukleinsäuren aneinander gebunden. Die Einstellung stringenter Bedingungen ist dem Fachmann bekannt und ist z.B. in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. beschrieben.
- 35
- c) Isolierung der kodierenden metY-Gene

Die für das Enzym O-Acetyl-Homoserin-Sulphydrolase kodierenden metY Gene aus den Organismen obiger Liste I sind in an sich bekannter Weise isolierbar.

5 Zur Isolierung der metY-Gene oder auch anderer Gene der Organismen aus obiger Liste I wird zunächst eine Genbank dieses Organismus in Escherichia coli (E. coli) angelegt. Das Anlegen von Genbanken ist in allgemein bekannten Lehrbüchern und Handbüchern ausführlich beschrieben. Als Beispiel seien das Lehrbuch von Winnacker: Gene und Klone, Eine Einführung in die Gentechnologie (Verlag Chemie, Weinheim, Deutschland, 1990), oder das Handbuch von Sambrook et al.: Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, 10 1989) genannt. Eine sehr bekannte Genbank ist die des E. coli K-12 Stammes W3110, die von Kohara et al. (Cell 50, 495-508 (198)) in λ -Vektoren angelegt wurde.

15 Zur Herstellung einer Genbank von Organismen der Liste I in E. coli können Cosmide, wie der Cosmidvektor SuperCos I (Wahl et al., 1987, Proceedings of the National Academy of Sciences USA, 84: 2160-2164), aber auch Plasmide, wie pBR322 (BollVal; Life Sciences, 25, 807-818 (1979)) oder pUC9 (Vieira et al., 1982, Gene, 19: 259-268), verwendet werden. Als Wirte eignen sich besonders solche E. coli Stämme, die restriktions- und rekombinationsdefekt sind. Ein Beispiel hierfür ist der Stamm DH5 α mc^r, der von Grant et al. (Proceedings of the National Academy of Sciences USA, 87 (1990) 4645-4649) beschrieben wurde. Die mit Hilfe von Cosmiden klonierten langen DNA-Fragmente können anschließend wiederum in gängige, für die Sequenzierung 20 geeignete Vektoren subkloniert und anschließend sequenziert werden, so wie es z. B. bei Sanger et al. (proceedings of the National Academy of Sciences of the United States of America, 74: 5463-5467, 1977) beschrieben ist.

25 Die erhaltenen DNA-Sequenzen können dann mit bekannten Algorithmen bzw. Sequenzanalyse-Programmen, wie z. B. dem von Staden (Nucleic Acids Research 14, 217-232 (1986)), dem von Marck (Nucleic Acids Research 16, 1829-1836 (1988)) oder dem GCG-Programm von Butler (Methods of Biochemical Analysis 39, 74-97 (1998)), untersucht werden.

30 Die für die metY-Gene kodierenden DNA-Sequenzen von Organismen gemäß obiger Liste I wurden gefunden. Insbesondere wurden DNA-Sequenzen gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 gefunden. Weiterhin wurde aus diesen vorliegenden DNA-Sequenzen mit den oben beschriebenen Methoden die Aminosäuresequenzen der entsprechenden Proteine abgeleitet. Durch SEQ ID NO: 2, 4, 35 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 und 54 sind die sich ergebenden Aminosäuresequenzen der metY Genprodukte dargestellt.

Kodierende DNA-Sequenzen, die sich aus den Sequenzen gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 durch die Degeneration des genetischen Kodes ergeben, sind ebenfalls Gegenstand der Erfindung. In gleicher Weise sind DNA-Sequenzen, die mit diesen Sequenzen oder davon abgeleiteten Sequenzteilen hybridisieren, Gegenstand der Erfindung.

Anleitungen zur Identifizierung von DNA-Sequenzen mittels Hybridisierung findet der Fachmann unter anderem im Handbuch "The DIG System Users Guide für Filter Hybridization" der Firma Boehringer Mannheim GmbH (Mannheim, Deutschland, 1993) und bei Liebl et al. (International Journal of Systematic Bacteriology (1991) 41: 255-260). Anleitungen zur Amplifikation von DNA-Sequenzen mit Hilfe der Polymerase-Kettenreaktion (PCR) findet der Fachmann unter anderem im Handbuch von Gait: Oligonukleotide synthesis: A Practical Approach (IRL Press, Oxford, UK, 1984) und bei Newton und Graham: PCR (Spektrum Akademischer Verlag, Heidelberg, Deutschland, 1994).

Weiterhin ist bekannt, dass Änderungen am N- und/oder C- Terminus eines Proteins dessen Funktion nicht wesentlich beeinträchtigen oder sogar stabilisieren können. Angaben hierzu findet der Fachmann unter anderem bei Ben-Bassat et al. (Journal of Bacteriology 169: 751-757 (1987)), bei O'Regan et al. (Gene 77: 237-251 (1989)), bei Sahin-Toth et al. (Protein Sciences 3: 240-247 (1994)), bei Hochuli et al. (Bionotechnology 6: 1321-1325 (1988)) und in bekannten Lehrbüchern der Genetik und Molekularbiologie.

Aminosäuresequenzen, die sich in entsprechender Weise aus den SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 und 54 ergeben, sind ebenfalls Bestandteil der Erfindung.

d) Erfindungsgemäß verwendete Wirtszellen

Weitere Gegenstände der Erfindung betreffen als Wirtszelle dienende Mikroorganismen, insbesondere coryneforme Bakterien, die einen Vektor, insbesondere Pendelvektor oder Plasmidvektor, der wenigstens ein metY Gen erfindungsgemäßer Definition trägt, enthalten oder in denen ein erfindungsgemäßes metY Gen exprimiert bzw. verstärkt ist.

Diese Mikroorganismen können schwefelhaltige Feinchemikalien, insbesondere L-Methionin, aus Glucose, Saccharose, Lactose, Fructose, Maltose, Melasse, Stärke, Cellulose oder aus Glycerin und Ethanol herstellen. Vorzugsweise sind dies coryneforme Bakterien, insbesondere der

Gattung *Corynebacterium*. Aus der Gattung *Corynebacterium* ist insbesondere die Art *Corynebacterium glutamicum* zu nennen, die in der Fachwelt für ihre Fähigkeit bekannt ist, L-Aminosäuren zu produzieren.

- 5 Als Beispiele für geeignete Stämme coryneformer Bakterien sind solche der Gattung *Corynebacterium*, insbesondere der Art *Corynebacterium glutamicum* (*C. glutamicum*), wie *Corynebacterium glutamicum* ATCC 13032, *Corynebacterium acetoglutamicum* ATCC 15806, *Corynebacterium acetoacidophilum* ATCC 13870,
- 10 *Corynebacterium thermoaminogenes* FERM BP-1539, *Corynebacterium melassecola* ATCC 17965

oder

- 15 der Gattung *Brevibacterium*, wie *Brevibacterium flavum* ATCC 14067 *Brevibacterium lactofermentum* ATCC 13869 und *Brevibacterium divaricatum* ATCC 14020 zu nennen;
- oder davon abgeleitete Stämme, wie *Corynebacterium glutamicum* KFCC10065
- 20 *Corynebacterium glutamicum* ATCC21608

welche ebenfalls die gewünschte Feinchemikalie oder deren Vorstufe(n) produzieren.

Mit der Abkürzung KFCC ist die Korean Federation of Culture Collection gemeint, mit der Abkürzung ATCC die American type strain culture collection und mit der Abkürzung FERM die Sammlung des National institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Japan.

- 25

e) Durchführung der erfindungsgemäßen Fermentation

- 30 Erfindungsgemäß wurde festgestellt, dass coryneforme Bakterien nach Überexpression eines *metY* Gens aus Organismen der Liste I in vorteilhafter Weise schwefelhaltige Feinchemikalien, insbesondere L-Methionin, produzieren.

- 35 Zur Erzielung einer Überexpression kann der Fachmann unterschiedliche Maßnahmen einzeln oder in Kombination ergreifen. So kann die Kopienzahl der entsprechenden Gene erhöht werden, oder es kann die Promotor- und Regulationsregion oder die Ribosomenbindungsstelle, die sich stromaufwärts des Strukturgens befindet, mutiert werden. In gleicher Weise wirken Expres-

sionskassetten, die stromaufwärts des Strukturgens eingebaut werden. Durch induzierbare Promotoren ist es zusätzlich möglich, die Expression im Verlaufe der fermentativen L-Methionin-Produktion zu steigern. Durch Maßnahmen zur Verlängerung der Lebensdauer der mRNA wird ebenfalls die Expression verbessert. Weiterhin wird durch Verhinderung des Abbaus des Enzymproteins ebenfalls die Enzymaktivität verstärkt. Die Gene oder Genkonstrukte können entweder in Plasmiden mit unterschiedlicher Kopienzahl vorliegen oder im Chromosom integriert und amplifiziert sein. Alternativ kann weiterhin eine Überexpression der betreffenden Gene durch Veränderung der Medienzusammensetzung und Kulturführung erreicht werden.

Anleitungen hierzu findet der Fachmann unter anderem bei Martin et al. (Bionotechnology 5, 137-146 (1987)), bei Guerrero et al. (Gene 138, 35-41 (1994)), Tsuchiya und Morinaga (Bio/Technology 6, 428-430 (1988)), bei Eikmanns et al. (Gene 102, 93-98 (1991)), in der Europäischen Patentschrift 0472869, im US Patent 4,601,893, bei Schwarzer und Pühler (Biotechnology 9, 84-87 (1991)), bei Remscheid et al. (Applied and Environmental Microbiology 60, 126-132 (1994)), bei LaBarre et al. (Journal of Bacteriology 175, 1001-1007 (1993)), in der Patentanmeldung WO 96/15246, bei Malumbres et al. (Gene 134, 15-24 (1993)), in der japanischen Offenlegungsschrift JP-A-10-229891, bei Jensen und Hammer (Biotechnology and Bioengineering 58, 191-195 (1998)), bei Makrides (Microbiological Reviews 60 : 512-538 (1996)) und in bekannten Lehrbüchern der Genetik und Molekularbiologie.

Gegenstand der Erfindung sind deshalb auch Expressionskonstrukte, enthaltend unter der genetischen Kontrolle regulativer Nukleinsäuresequenzen eine für ein erfindungsgemäßes Polypeptid kodierende Nukleinsäuresequenz; sowie Vektoren, umfassend wenigstens eines dieser Expressionskonstrukte. Vorzugsweise umfassen solche erfindungsgemäßen Konstrukte 5'-stromaufwärts von der jeweiligen kodierenden Sequenz einen Promotor und 3'-stromabwärts eine Terminatorsequenz sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils operativ verknüpft mit der kodierenden Sequenz. Unter einer „operativen Verknüpfung“ versteht man die sequentielle Anordnung von Promotor, kodierender Sequenz, Terminator und gegebenenfalls weiterer regulativer Elemente derart, dass jedes der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann. Beispiele für operativ verknüpfbare Sequenzen sind Aktivierungssequenzen sowie Enhancer und dergleichen. Weitere regulative Elemente umfassen selektierbare Marker, Amplifikationssignale, Replikationsursprünge und dergleichen. Geeignete regulatorische Sequenzen sind z.B. beschrieben in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990).

Zusätzlich zu den artifiziellen Regulationssequenzen kann die natürliche Regulationssequenz

vor dem eigentlichen Strukturgen noch vorhanden sein. Durch genetische Veränderung kann diese natürliche Regulation gegebenenfalls ausgeschaltet und die Expression der Gene erhöht oder erniedrigt werden. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das heißt es werden keine zusätzlichen Regulationssignale vor das Strukturgen insertiert und der natürliche Promotor mit seiner Regulation wird nicht entfernt. Statt dessen wird die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert oder verringert wird. Die Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

- 10 Beispiele für brauchbare Promotoren sind: die Promotoren, *ddh*, *amy*, *lysC*, *dapA*, *lysA* aus *Corynebacterium glutamicum*, aber auch gram-positiven Promotoren SPO2 wie sie in *Bacillus Subtilis* and Its Closest Relatives, Sonenshein, Abraham L., Hoch, James A., Losick, Richard; ASM Press, District of Columbia, Washington und Patek M. Eikmanns B.J. Patek J. Sahm H. Microbiology. 142 1297-309, 1996 beschrieben sind, oder aber auch *cos*-, *tac*-, *trp*-, *tet*-, *trp-tet*-, *lpp*-, *lac*-, *lpp-lac*-, *lacIq*-, *T7*-, *T5*-, *T3*-, *gal*-, *trc*-, *ara*-, *SP6*-, λ -PR- oder im λ -PL-Promotor, die vorteil-
- 15 , *lpp-lac*-, *lacIq*-, *T7*-, *T5*-, *T3*-, *gal*-, *trc*-, *ara*-, *SP6*-, λ -PR- oder im λ -PL-Promotor, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden. Bevorzugt ist auch die Verwendung induzierbarer Promotoren, wie z.B. licht- und insbesondere temperaturinduzierbarer Promotoren, wie der P_{P_1} -Promotor. Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen verwendet werden. Darüber hinaus können auch synthetische Promotoren vor-
- 20 teilhaft verwendet werden.

Die genannten regulatorischen Sequenzen sollen die gezielte Expression der Nukleinsäuresequenzen ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird.

- 25 Die regulatorischen Sequenzen bzw. Faktoren können dabei vorzugsweise die Expression positiv beeinflussen und dadurch erhöhen oder erniedrigen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.
- 30

Die Herstellung einer Expressionskassette erfolgt durch Fusion eines geeigneten Promotors, einer geeigneten Shine-Dalgarno-Sequenz mit einer metY-Nukleotidsequenz sowie einem geeigneten Terminationssignal. Dazu verwendet man gängige Rekombinations- und Klonierungstechniken, wie sie beispielsweise in Current Protocols in Molecular Biology, 1993, John

35

- Wiley & Sons, Incorporated, New York New York, PCR Methods, Gelfand, David H., Innis, Michael A., Sninsky, John J. 1999, Academic Press, Incorporated, California, San Diego, ., PCR Cloning Protocols, Methods in Molecular Biology Ser., Vol. 192, 2nd ed., Humana Press, New Jersey, Totowa. T. Maniatis, E.F. Fritsch und J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience (1987) beschrieben sind.
- 10 Das rekombinante Nukleinsäurekonstrukt bzw. Genkonstrukt wird zur Expression in einem geeigneten Wirtsorganismus vorteilhafterweise in einen wirtsspezifischen Vektor insertiert, der eine optimale Expression der Gene im Wirt ermöglicht. Vektoren sind dem Fachmann wohl bekannt und können beispielsweise aus "Cloning Vectors" (Pouwels P. H. et al., Hrsg, Elsevier, Amsterdam-New York-Oxford, 1985) entnommen werden. Unter Vektoren sind außer Plasmiden auch
- 15 alle anderen dem Fachmann bekannten Vektoren, wie beispielsweise Phagen, Transposons, IS-Elemente, Phasmide, Cosmide, und lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können autonom im Wirtsorganismus repliziert oder chromosomal repliziert werden.
- 20 Zur Verstärkung wurden erfindungsgemäße metY Gene beispielhaft mit Hilfe von episomalen Plasmiden überexprimiert. Als Plasmide eignen sich solche, die in coryneformen Bakterien repliziert werden. Zahlreiche bekannte Plasmidvektoren, wie z. B. pZ1 (Menkel et al., Applied and Environmental Microbiology (1989) 64: 549-554), pEKEx1 (Eikmanns et al., Gene 102: 93-98 (1991)) oder pHS2-1 (Sonnen et al., Gene 107: 69-74 (1991)) beruhen auf den kryptischen Plasmiden pHM1519, pBL1 oder pGA1. Andere Plasmidvektoren, wie z. B. pCLiK5MCS, oder
- 25 solche, die auf pCG4 (US-A 4,489,160) oder pNG2 (Serwold-Davis et al., FEMS Microbiology Letters 66, 119-124 (1990)) oder pAG1 (US-A 5,158,891) beruhen, können in gleicher Weise verwendet werden.
- 30 Weiterhin eignen sich auch solche Plasmidvektoren mit Hilfe derer man das Verfahren der Genamplifikation durch Integration in das Chromosom anwenden kann, so wie es beispielsweise von Remscheid et al. (Applied and Environmental Microbiology 60,126-132 (1994)) zur Duplikation bzw. Amplifikation des hom-thrB-Operons beschrieben wurde. Bei dieser Methode wird das vollständige Gen in einen Plasmidvektor kloniert, der in einem Wirt (typischerweise E. coli), nicht aber in C. glutamicum replizieren kann. Als Vektoren kommen beispielsweise pSUP301 (Simon
- 35 et al., Bio/ Technology 1,784-791 (1983)), pK18mob oder pK19mob (Schäfer et al., Gene 145,69-73 (1994)), Bernard et al., Journal of Molecular Biology, 234: 534-541 (1993)), pEM1 (Schrumpf et al. 1991, Journal of Bacteriology 173: 4510-4516) oder pBGS8 (Spratt et al., 1986,

Gene 41: 337-342) in Frage. Der Plasmidvektor, der das zu amplifizierende Gen enthält, wird anschließend durch Transformation in den gewünschten Stamm von *C. glutamicum* überführt. Methoden zur Transformation sind beispielsweise bei Thierbach et al. (Applied Microbiology and Biotechnology 29, 356-362 (1988)), Dunican und Shivnan (Biotechnology 7, 1067-1070 (1989)) und Tauch et al. (FEMS Microbiological Letters 123,343-347 (1994)) beschrieben.

Enzyme können durch Mutationen in den korrespondierenden Genen derart in ihrer Aktivität beeinflusst werden, dass es zu einer teilweisen oder vollständigen Verringerung der Reaktionsgeschwindigkeit der enzymatischen Reaktion kommt. Beispiele für solche Mutationen sind dem Fachmann bekannt (Motoyama H. Yano H. Terasaki Y. Anazawa H. Applied & Environmental Microbiology. 67:3064-70, 2001, Eikmanns B.J. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94.)

Zusätzlich kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben einer Expression bzw. Verstärkung eines erfindungsgemäßen metY-Gen eines oder mehrere Enzyme des jeweiligen Biosyntheseweges, des Cystein-Stoffwechselwegs, der Aspartatsemialdehyd-Synthese, der Glykolyse, der Anaplerotik, des Pentose-Phosphat-Stoffwechsels, des Zitronensäure-Zyklus oder des Aminosäure-Exports zu verstärken.

So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, eines oder mehrere der folgenden Gene verstärkt sein:

- das für eine Aspartatkinase kodierende Gen *lysC* (EP 1 108 790 A2; DNA-SEQ NO. 281),
- das für eine Aspartatsemialdehyd Dehydrogenase kodierende Gen *asd* (EP 1 108 790 A2; DNA-SEQ NO. 282),
- das für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierende Gen *gap* (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die 3-Phosphoglycerat Kinase kodierende Gen *pgk* (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die Pyruvat Carboxylase kodierende Gen *pyc* (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die Triosephosphat Isomerase kodierende Gen *tpi* (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- das für die Homoserin O-Acetyltransferase kodierende Gen *metA* (EP 1 108 790 A2; DNA-SEQ NO. 725),
- das für die Cystathionin-gamma-Synthase kodierende Gen *metB* (EP 1 108 790 A2; DNA-SEQ NO. 3491),

- das für die Cystathionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),
- das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ NO. 1110),
- 5 - das für die MethioninSynthase kodierende Gen meth (EP 1 108 790 A2),
- das für die Methylentetrahydrofolat-Reduktase kodierende Gen metF (EP 1 108 790 A2; DNA-SEQ NO. 2379),
- das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-SEQ NO. 928)
- 10 - eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)
- das für die Serine Acetyl-Transferase kodierende Gen cysE (EP 1 108 790 A2; DNA-SEQ NO. 2818)
- das für eine Homoserin-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ
- 15 NO. 1306)

So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, in coryneformen Bakterien, vorteilhaft sein, gleichzeitig wenigstens eines der nachfolgenden Gene zu mutieren, so dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch einen Stoffwechselmetaboliten in ihrer Aktivität beeinflusst werden:

- das für eine Aspartatkinase kodierende Gen lysC (EP 1 108 790 A2; DNA-SEQ NO. 281),
- das für die Pyruvat Carboxylase kodierende Gen pyc (Eikmanns (1992), Journal of Bacteriology
- 25 174: 6076-6086),
- das für die Homoserin O-Acetyltransferase kodierende Gen metA (EP 1 108 790 A2; DNA-SEQ NO. 725),
- das für die Cystathionin-gamma-Synthase kodierende Gen metB (EP 1 108 790 A2; DNA-SEQ NO. 3491),
- 30 - das für die Cystathionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),
- das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ NO. 1110),
- das für die Methionin Synthase kodierende Gen meth (EP 1 108 790 A2),
- 35 - das für die Methylentetrahydrofolat-Reduktase kodierende Gen metF (EP 1 108 790 A2; DNA-SEQ NO. 2379),
- das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-

SEQ NO. 928)

- eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)

- das für die Serine Acetyl-Transferase kodierende Gen cysE (EP 1 108 790 A2; DNA-SEQ NO.

5 2818)

- das für eine Homoserin-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ NO. 1306)

10 Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsge-
mäßigen metY-Gene eines oder mehrere der folgenden Gene abzuschwächen, insbesondere de-
ren Expression zu verringern, oder auszuschalten:

- das für die Homoserine-Kinase kodierende Gen thrB (EP 1 108 790 A2; DNA-SEQ NO. 3453)

15 - das für die Threonin Dehydratase kodierende Gen ilvA (EP 1 108 790 A2; DNA-SEQ NO. 2328)

- das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)

- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2; DNA-SEQ NO. 3494)

20 - das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)

- das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)

- das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)

25 - das für die Dihydrodipicolinat Synthase kodierende Gen dapA (EP 1 108 790 A2; DNA-SEQ NO. 3476)

- das für die Dihydrodipicolinat Reduktase kodierende Gen dapB (EP 1 108 790 A2; DNA-SEQ NO. 3477)

30 - das für die Diaminopicolinat Decarboxylase kodierende Gen lysA (EP 1 108 790 A2; DNA-SEQ NO. 3451)

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsge-
mäßigen metY-Gene in Coryneformen Bakterien gleichzeitig wenigstens eines der folgenden Ge-
ne so zu mutieren, dass die enzymatische Aktivität des korrespondierenden Proteins teilweise
35 oder vollständig verringert wird:

- das für die Homoserine-Kinase kodierende Gen thrB (EP 1 108 790 A2; DNA-SEQ NO. 3453)
- das für die Threonin Dehydratase kodierende Gen ilvA (EP 1 108 790 A2; DNA-SEQ NO. 2328)
- das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)
- 5 - das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2; DNA-SEQ NO. 3494)
- das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)
- 10 - das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)
- das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)
- das für die Dihydrodipicolinat Synthase kodierende Gen dapA (EP 1 108 790 A2; DNA-SEQ NO. 3476)
- 15 - das für die Dihydrodipicolinat Reduktase kodierende Gen dapB (EP 1 108 790 A2; DNA-SEQ NO. 3477)
- das für die Diaminopicolinat Decarboxylase kodierende Gen lysA (EP 1 108 790 A2; DNA-SEQ NO. 3451)

20 Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben der Expression bzw. Verstärkung eines erfindungsgemäßen metY-Gens unerwünschte Nebenreaktionen auszuschalten (Nakayama: "Breeding of Amino Acid Producing Microorganisms", in: Overproduction of Microbial Products, Krumphanzl, Sikyta, Vanek (eds.), Academic Press, London, UK, 1982).

25 Die erfindungsgemäß hergestellten Mikroorganismen können kontinuierlich oder diskontinuierlich im batch-Verfahren (Satzkultivierung) oder im fed batch (Zulaufverfahren) oder repeated fed batch Verfahren (repetitives Zulaufverfahren) zur Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, kultiviert werden. Eine Zusammenfassung über bekannte Kultivierungsmethoden ist im Lehrbuch von Chmiel (Bioprozeßtechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) oder im Lehrbuch von Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)) zu finden.

30

Das zu verwendende Kulturmedium hat in geeigneter Weise den Ansprüchen der jeweiligen Stämme zu genügen. Beschreibungen von Kulturmedien verschiedener Mikroorganismen sind

35 im Handbuch "Manual of Methods für General Bacteriology" der American Society für Bacteriology (Washington D. C., USA, 1981) enthalten.

Diese erfindungsgemäß einsetzbaren Medien umfassen gewöhnlich eine oder mehreren Kohlenstoffquellen, Stickstoffquellen, anorganische Salze, Vitamine und/oder Spurenelemente.

5 Bevorzugte Kohlenstoffquellen sind Zucker, wie Mono-, Di- oder Polysaccharide. Sehr gute Kohlenstoffquellen sind beispielsweise Glucose, Fructose, Mannose, Galactose, Ribose, Sorbose, Ribulose, Lactose, Maltose, Saccharose, Raffinose, Stärke oder Cellulose. Man kann Zucker auch über komplexe Verbindungen, wie Melassen, oder andere Nebenprodukte der Zucker-Raffinierung zu den Medien geben. Es kann auch vorteilhaft sein, Gemische verschiedener Kohlenstoffquellen zuzugeben. Andere mögliche Kohlenstoffquellen sind Öle und Fette wie z. B. Sojaöl. Sonnenblumenöl. Erdnußöl und Kokosfett, Fettsäuren wie z. B. Palmitinsäure, Stearinsäure oder Linolsäure, Alkohole wie z. B. Glycerin, Methanol oder Ethanol und organische Säuren wie z. B. Essigsäure oder Milchsäure.

15 Stickstoffquellen sind gewöhnlich organische oder anorganische Stickstoffverbindungen oder Materialien, die diese Verbindungen enthalten. Beispielhafte Stickstoffquellen umfassen Ammoniak-Gas oder Ammoniumsalze, wie Ammoniumsulfat, Ammoniumchlorid, Ammoniumphosphat, Ammoniumcarbonat oder Ammoniumnitrat, Nitrate, Harnstoff, Aminosäuren oder komplexe Stickstoffquellen, wie Maisquellwasser, Sojamehl, Sojaprotein, Hefeextrakt, Fleischextrakt und andere. Die Stickstoffquellen können einzeln oder als Mischung verwendet werden.

20 Anorganische Salzverbindungen, die in den Medien enthalten sein können, umfassen die Chlorid-, Phosphor- oder Sulfatsalze von Calcium, Magnesium, Natrium, Kobalt, Molybdän, Kalium, Mangan, Zink, Kupfer und Eisen

25 Als Schwefelquelle für die Herstellung von schwefelhaltigen Feinchemikalien, insbesondere von Methionin, können anorganische schwefelhaltige Verbindungen wie beispielsweise Sulfate, Sulfite, Dithionite, Tetrathionate, Thiosulfate, Sulfide aber auch organische Schwefelverbindungen, wie Mercaptane und Thiole, verwendet werden.

30 Als Phosphorquelle können Phosphorsäure, Kaliumdihydrogenphosphat oder Dikaliumhydrogenphosphat oder die entsprechenden Natrium haltigen Salze verwendet werden.

35 Chelatbildner können zum Medium gegeben werden, um die Metallionen in Lösung zu halten. Besonders geeignete Chelatbildner umfassen Dihydroxyphenole, wie Catechol oder Protocatechuat, oder organische Säuren, wie Citronensäure.

Die erfindungsgemäß eingesetzten Fermentationsmedien enthalten üblicherweise auch andere

Wachstumsfaktoren, wie Vitamine oder Wachstumsförderer, zu denen beispielsweise Biotin, Riboflavin, Thiamin, Folsäure, Nikotinsäure, Panthothenat und Pyridoxin gehören. Wachstumsfaktoren und Salze stammen häufig von komplexen Medienkomponenten, wie Hefeextrakt, Melassen, Maisquellwasser und dergleichen. Dem Kulturmedium können überdies geeignete Vorstufen zugesetzt werden. Die genaue Zusammensetzung der Medienverbindungen hängt stark vom jeweiligen Experiment ab und wird für jeden spezifischen Fall individuell entschieden. Information über die Medienoptimierung ist erhältlich aus dem Lehrbuch "Applied Microbiol. Physiology, A Practical Approach" (Hrsg. P.M. Rhodes, P.F. Stanbury, IRL Press (1997) S. 53-73, ISBN 0 19 963577 3). Wachstumsmedien lassen sich auch von kommerziellen Anbietern beziehen, wie Standard 1 (Merck) oder BHI (Brain heart infusion, DIFCO) und dergleichen.

Sämtliche Medienkomponenten werden, entweder durch Hitze (20 min bei 1,5 bar und 121°C) oder durch Sterilfiltration, sterilisiert. Die Komponenten können entweder zusammen oder nötigenfalls getrennt sterilisiert werden. Sämtliche Medienkomponenten können zu Beginn der Anzucht zugegen sein oder wahlfrei kontinuierlich oder chargenweise hinzugegeben werden.

Die Temperatur der Kultur liegt normalerweise zwischen 15°C und 45°C, vorzugsweise bei 25°C bis 40°C und kann während des Experimentes konstant gehalten oder verändert werden. Der pH-Wert des Mediums sollte im Bereich von 5 bis 8,5, vorzugsweise um 7,0 liegen. Der pH-Wert für die Anzucht läßt sich während der Anzucht durch Zugabe von basische Verbindungen wie Natriumhydroxid, Kaliumhydroxid, Ammoniak bzw. Ammoniakwasser oder saure Verbindungen wie Phosphorsäure oder Schwefelsäure kontrollieren. Zur Kontrolle der Schaumentwicklung können Antischaummittel wie z. B. Fettsäurepolyglykolester, eingesetzt werden. Zur Aufrechterhaltung der Stabilität von Plasmiden können dem Medium geeignete selektiv wirkende Stoffe, wie z. B. Antibiotika, hinzugefügt werden. Um aerobe Bedingungen aufrechtzuerhalten, werden Sauerstoff oder Sauerstoff haltige Gasmischungen, wie z. B. Umgebungsluft, in die Kultur eingetragen. Die Temperatur der Kultur liegt normalerweise bei 20°C bis 45°C. Die Kultur wird solange fortgesetzt, bis sich ein Maximum des gewünschten Produktes gebildet hat. Dieses Ziel wird normalerweise innerhalb von 10 Stunden bis 160 Stunden erreicht.

Die so erhaltenen, insbesondere L-Methionin enthaltenden, Fermentationsbrühen haben üblicherweise eine Trockenmasse von 7,5 bis 25 Gew.-%.

Vorteilhaft ist außerdem auch, wenn die Fermentation zumindest am Ende, insbesondere jedoch über mindestens 30% der Fermentationsdauer zuckerlimitiert gefahren wird. Das heißt, dass während dieser Zeit die Konzentration an verwertbarem Zucker im Fermentationsmedium auf ≥ 0 bis 3 g/l gehalten, beziehungsweise abgesenkt wird.

Die Fermentationsbrühe wird anschließend weiterverarbeitet. Je nach Anforderung kann die Biomasse ganz oder teilweise durch Separationsmethoden, wie z. B. Zentrifugation, Filtration, Dekantieren oder einer Kombination dieser Methoden aus der Fermentationsbrühe entfernt oder vollständig in ihr belassen werden.

Anschließend kann die Fermentationsbrühe mit bekannten Methoden, wie z. B. mit Hilfe eines Rotationsverdampfers, Dünnschichtverdampfers, Fallfilmverdampfers, durch Umkehrosmose, oder durch Nanofiltration, eingedickt beziehungsweise aufkonzentriert werden. Diese aufkonzentrierte Fermentationsbrühe kann anschließend durch Gefriertrocknung, Sprühtrocknung, Sprühgranulation oder durch anderweitige Verfahren aufgearbeitet werden.

Es ist aber auch möglich die schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, weiter aufzureinigen. Hierzu wird die produktthaltige Brühe nach dem Abtrennen der Biomasse einer Chromatographie mit einem geeigneten Harz unterworfen, wobei das gewünschte Produkt oder die Verunreinigungen ganz oder teilweise auf dem Chromatographieharz zurückgehalten werden. Diese Chromatographieschritte können nötigenfalls wiederholt werden, wobei die gleichen oder andere Chromatographieharze verwendet werden. Der Fachmann ist in der Auswahl der geeigneten Chromatographieharze und ihrer wirksamsten Anwendung bewandert. Das gereinigte Produkt kann durch Filtration oder Ultrafiltration konzentriert und bei einer Temperatur aufbewahrt werden, bei der die Stabilität des Produktes maximal ist.

Die Identität und Reinheit der isolierten Verbindung(en) kann durch Techniken des Standes der Technik bestimmt werden. Diese umfassen Hochleistungs-Flüssigkeitschromatographie (HPLC), spektroskopische Verfahren, Färbeverfahren, Dünnschichtchromatographie, NIRS, Enzymtest oder mikrobiologische Tests. Diese Analyseverfahren sind zusammengefaßt in: Patek et al. (1994) Appl. Environ. Microbiol. 60:133-140; Malakhova et al. (1996) Biotekhnologiya 11 27-32; und Schmidt et al. (1998) Bioprocess Engineer. 19:67-70. Ulmann's Encyclopedia of Industrial Chemistry (1996) Bd. A27, VCH: Weinheim, S. 89-90, S. 521-540, S. 540-547, S. 559-566, 575-581 und S. 581-587; Michal, G (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. et al. (1987) Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17.

Die Erfindung wird nun anhand der folgenden nicht-limitierenden Beispiele und unter Bezugnahme auf beiliegende Figuren näher beschrieben. Dabei zeigt

Figur 1 die Plasmidkarte zu Plasmid pClysC;

Figur 2 die Plasmidkarte zu Plasmid pCISlysCthr311ile;

Figur 3 die Plasmidkarte zu Plasmid pCPhsdhmetY_Mt.

Restriktionsschnittstellen mit der entsprechenden Positionsangabe in Klammern sind in den Plasmidkarten angegeben. Wesentliche Sequenzabschnitte sind fettgedruckt beschrieben.

5 KanR steht für Kanamycin-Resistenzgen; ask steht für Aspartatkinasegen

Beispiel 1: Konstruktion von pCLiK5MCS

10 Zunächst wurden Ampicillinresistenz und Replikationsursprung des Vektors pBR322 mit den Oligonukleotiden p1.3 (SEQ ID NO:55) und p2.3 (SEQ ID NO:56) mit Hilfe der Polymerase-Kettenreaktion (PCR) amplifiziert.

p1.3 (SEQ ID NO:55)

15 5'-**CCCGGGATCCGCTAGCGGCGCGCCGGCCGGCCCGGTGTGAAATACCGCACAG**-3'

p2.3 (SEQ ID NO:56)

5'-**TCTAGACTCGAGCGGCCGCGGCCGGCCTTTAAATTGAAGACGAAAGGGCCTCG**-3'

20 Neben den zu pBR322 komplementären Sequenzen, enthält das Oligonukleotid p1.3 (SEQ ID NO:55) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen SmaI, BamHI, NheI und AscI und das Oligonukleotid p2.3 (SEQ ID NO:56) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen XbaI, XhoI, NotI und DraI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,1 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Die stumpfen Enden des DNA-Fragmentes wurden mit dem Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers miteinander ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

35 Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK1.

Ausgehend vom Plasmid pWLT1 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden neo1 (SEQ ID NO:57) und neo2 (SEQ ID NO:58) eine Kanamycin-Resistenzcassette amplifiziert.

5

neo1 (SEQ ID NO:57):

5'-GAGATCTAGACCCGGGGATCCGCTAGCGGGCTGCTAAAGGAAGCGGA-3'

neo2 (SEQ ID NO:58):

10 5'-GAGAGGCGCGCCGCTAGCGTGGGCGAAGAACTCCAGCA-3'

15

20

25

30

Neben den zu pWLT1 komplementären Sequenzen, enthält das Oligonukleotid neo1 in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen XbaI, SmaI, BamHI, NheI und das Oligonukleotid neo2 (SEQ ID NO:58) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen AscI und NheI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,3 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit den Restriktionsendonukleasen XbaI und AscI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK1 wurde ebenfalls mit den Restriktionsendonukleasen XbaI und AscI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,1kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) und Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

35

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so

erhaltene Plasmid erhält den Namen pCLiK2.

Der Vektor pCLiK2 wurde mit der Restriktionsendonuklease DraI (New England Biolabs, Beverly, USA) geschnitten. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde ein ca. 2,3 kb
5 großes Vektorfragment mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers religiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben (1989)), in kompetente E.coli XL-
10 1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so
15 erhaltene Plasmid erhält den Namen pCLiK3.

Ausgehend vom Plasmid pWLQ2 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden cg1 ((SEQ ID NO:59) und cg2 (SEQ ID NO:60) der Replikationsursprung pHM1519 amplifiziert.
20

cg1 (SEQ ID NO:59):

5'-GAGAGGGCGGCCGCGCAAAGTCCCGCTTCGTGAA-3'

25 cg2 (SEQ ID NO:60):

5'-GAGAGGGCGGCCGCTCAAGTCGGTCAAGCCACGC-3'

Neben den zu pWLQ2 komplementären Sequenzen, enthalten die Oligonukleotide cg1 (SEQ ID NO:59) und cg2 (SEQ ID NO:60) Schnittstellen für die Restriktionsendonuklease NotI. Die PCR
30 Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,7 kb wurde mit dem
GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease NotI
35 (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben

des Herstellers gereinigt. Der Vektor pCLiK3 wurde ebenfalls mit der Restriktionsendonuklease NotI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,3kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5.

Für die Erweiterung von pCLiK5 um eine „multiple cloning site“ (MCS) wurden die beide synthetischen, weitestgehend komplementären Oligonukleotide HS445 ((SEQ ID NO:61) und HS446 ((SEQ ID NO:62), die Schnittstellen für die Restriktionsendonukleasen SwaI, XhoI, AatI, ApaI, Asp718, MluI, NdeI, SpeI, EcoRV, Sall, ClaI, BamHI, XbaI und SmaI enthalten, durch gemeinsames erhitzen auf 95°C und langsames abkühlen zu einem doppelsträngigen DNA-Fragment vereinigt.

HS445 (SEQ ID NO:61):
5'-TCGAATTAAATCTCGAGAGGCCTGACGTCGGGCCCGGTACCACGCGTCATATGACTAG
TTCGGACCTAGGGATATCGTCGACATCGATGCTCTTCTGCGTTAATTAACAATTGGGATCC
TCTAGACCCGGGATTAAAT-3'

HS446 (SEQ ID NO:62):
5'-GATCATTTAAATCCCGGGTCTAGAGGATCCCAATTGTTAATTAACGCAGAAGAGCATCGA
TGTCGACGATATCCCTAGGTCCGAAGTAGTCATATGACGCGTGGTACCGGGCCCGACGTC
AGGCCTCTCGAGATTAAAT-3'

Der Vektor pCLiK5 wurde mit den Restriktionsendonuklease XhoI und BamHI (New England Biolabs, Beverly, USA) geschnitten und mit alkalischer Phosphatase (I (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem

0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 5,0 kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem synthetischen Doppelsträngigen DNA-Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5MCS.

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierungsreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pCLiK5MCS ist als SEQ ID NO: 65 aufgeführt.

Beispiel 2: Konstruktion von pCLiK5MCS integrativ sacB

Ausgehend vom Plasmid pK19mob (Schäfer et al., Gene 145,69-73(1994)) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden BK1732 und BK1733 das Bacillus subtilis sacB Gen (kodierend für Levan Sucrase) amplifiziert.

BK1732 (SEQ ID NO:63):

5'-GAGAGCGGCCGCGATCCTTTTAAACCCATCAC-3'

BK1733 (SEQ ID NO:64):

5'-AGGAGCGGCCGCGCATCGGCATTTTCTTTTGCG-3'

Neben den zu pEK19mobsac komplementären Sequenzen, enthalten die Oligonukleotide BK1732 und BK1733 Schnittstellen für die Restriktionsendonuklease NotI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchge-

führt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,9 kb wurde mit dem
GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben
des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease NotI
(New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem
5 GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben
des Herstellers gereinigt.

Der Vektor pCLiK5MCS (hergestellt gemäß Beispiel 1) wurde ebenfalls mit der Restriktionsen-
donuklease NotI geschnitten und mit alkalischer Phosphatase (I (Roche Diagnostics, Mann-
10 heim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen
Agarosegel wurde ein ungefähr 2,4 kb großes Vektorfragment mit dem GFX™PCR, DNA and
Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers iso-
liert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics,
Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der
15 Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laborato-
ry Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene,
La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Aus-
plattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

20 Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen,
Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so
erhaltene Plasmid erhält den Namen pCLiK5MCS integrativ sacB.

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Aca-
25 demy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels
ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pCLiK5MCS integrativ sacB ist als SEQ ID NO: 66 aufgeführt.

30 Weitere Vektoren die zur erfindungsgemäßen Expression oder Überproduktion von metY-Genen
geeignet sind, können in analoger Weise hergestellt werden.

Beispiel 3: Isolierung des lysC Gens aus dem C. glutamicum Stamm LU1479

35 Im ersten Schritt der Stammkonstruktion soll ein allelischer Austausch des lysC Wildtypgens,
kodierend für das Enzym Aspartatkinase, in C. glutamicum ATCC13032, im folgenden LU1479

genannt, durchgeführt werden. Dabei soll im LysC Gen ein Nukleotidaustausch durchgeführt werden, so dass im resultierenden Protein die Aminosäure Thr an der Position 311 durch die Aminosäure Ile ausgetauscht ist.

- 5 Ausgehend von der chromosomalen DNA aus LU1479 als Template für eine PCR Reaktion wurde mit den Oligonukleotidprimern SEQ ID NO:67 und SEQ ID NO:68 lysC mit Hilfe des Pfu-Turbo PCR Systems (Stratagene USA) nach Angaben des Herstellers amplifiziert. Chromosomale DNA aus *C. glutamicum* ATCC 13032 wurde nach Tauch et al. (1995) Plasmid 33:168-179 oder Eikmanns et al. (1994) Microbiology 140:1817-1828 präpariert. Das amplifizierte Fragment
- 10 wird an seinem 5'-Ende von einem Sall Restriktionsschnitt und an seinem 3'-Ende von einem MluI Restriktionsschnitt flankiert. Vor der Klonierung wurde das amplifizierte Fragment durch diese beiden Restriktionsenzyme verdaut und mit GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aufgereinigt.

15 SEQ ID NO:67
5'-GAGAGAGAGACGCGTCCCAGTGGCTGAGACGCATC-3'

SEQ ID NO:68
5'-CTCTCTCTGTCGACGAATTCAATCTTACGGCCTG-3'

- 20 Das erhaltenen Polynukleotid wurde über die Sall und MluI Restriktionsschnitte in pCLIK5 MCS integrativ SacB (im folgenden pCIS genannt; SEQ ID NO: 66 aus Beispiel 2) kloniert und in *E.coli* XL-1 blue transformiert. Eine Selektion auf Plasmid-tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml)-haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht. Das Plasmid wurden isoliert und durch Sequenzierung die erwartete Nukleotidsequenz
- 25 bestätigt. Die Präparation der Plasmid-DNA wurde nach Methoden und mit Materialien der Firma Qiagen durchgeführt. Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierungsreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und
- 30 ausgewertet. Das erhaltene Plasmid pCIS lysC ist als SEQ ID NO:69 aufgeführt. Die entsprechende Plasmidkarte ist in Figur 1 dargestellt.

Die Sequenz SEQ ID NO:69 umfasst die folgenden wesentlichen Teilbereiche:

LOCUS pCIS\lysC 5860 bp DNA circular

35 FEATURES Location/Qualifiers
CDS¹⁾ 155..1420
/vntifkey="4"

5 /label=lysC
 complement²⁾(3935..5356)
 /vntifkey="4"
 /label=sacB\Bacillus\subtilis)
 10 promoter complement(5357..5819)
 /vntifkey="30"
 /label=Promotor\sacB
 complement(3913..3934)
 /vntifkey="2"
 /label=sacB\downstreambereich
 1974..2765
 /vntifkey="4"
 /label=Kan\R
 15 CDS complement(3032..3892)
 /vntifkey="4"
 /label=Ori\EC\pMB)

¹⁾ kodierende Sequenz

²⁾ auf Komplementärstrang

20 Beispiel 4: Mutagenese des lysC Gens aus *C. glutamicum*

Die gerichtete Mutagenese des lysC Gens aus *C. glutamicum* (Beispiel 3) wurde mit dem QuickChange Kit (Fa. Stratagene/USA) nach Angaben des Herstellers durchgeführt. Die Mutagenese wurde im Plasmid pCIS lysC, SEQ ID NO:69 durchgeführt. Für den Austausch von thr311 nach 311ile mit Hilfe der Quickchange Methode (Stratagene) wurden folgende Oligonukleotidprimer synthetisiert

SEQ ID NO:70

5'-CGGCACCACCGACATCATCTTCACCTGCCCTCGTTCCG -3'

30 SEQ ID NO:71

5'-CGGAACGAGGGCAGGTGAAGATGATGTCGGTGGTGCCG -3'

35 Der Einsatz dieser Oligonukleotidprimer in der Quickchange Reaktion führt in dem lysC Gen zu einem Austausch des Nukleotids in Position 932 (von C nach T) (vgl. SEQ ID NO:72) und im korrespondierenden Enzym zu einem Aminosäuresubstitution in Position 311 (Thr→Ile) (vgl. SEQ ID NO:73). Der resultierende Aminosäureaustausch Thr311Ile im lysC Gen wurde nach Transformation in *E.coli* XL1-blue und Plasmidpräparation durch Sequenzierung bestätigt. Das

Plasmid erhielt die Bezeichnung pCIS lysC thr311ile und ist als SEQ ID NO:74 aufgeführt. Die entsprechende Plasmidkarte ist in Figur 2 dargestellt.

5 Die Sequenz SEQ ID NO:74 umfasst die folgenden wesentlichen Teilbereiche:

	LOCUS	pCIS\lysC\thr311ile	5860 bp	DNA	circular
	FEATURES	Location/Qualifiers			
10	CDS ¹⁾	155..1420			
		/vntifkey="4"			
		/label=lysC			
	CDS	complement ²⁾ (3935..5356)			
		/vntifkey="4"			
		/label=sacB\Bacillus\subtilis			
15	promoter	complement(5357..5819)			
		/vntifkey="30"			
		/label=Promotor\sacB			
	C_region	complement(3913..3934)			
		/vntifkey="2"			
20		/label=sacB\downstreambereich			
	CDS	1974..2765			
		/vntifkey="4"			
		/label=Kan\R			
25	CDS	complement(3032..3892)			
		/vntifkey="4"			
		/label=Ori\EC\pMB			

¹⁾ kodierende Sequenz

²⁾ auf Komplementärstrang

30 Das Plasmid pCIS lysC thr311ile wurde in *C. glutamicum* LU1479 mittels Elektroporation wie bei Liebl, et al. (1989) FEMS Microbiology Letters 53:299-303 beschrieben, transformiert. Modifikationen des Protokolls sind in DE-A-10046870 beschrieben. Die chromosomale Anordnung des lysC-Lokus einzelner Transformanten wurde mit Standardmethoden durch Southernblot und Hybridisierung, wie in Sambrook et al. (1989), Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben, überprüft. Dadurch wurde sichergestellt, dass es sich bei den Transformanten um solche handelt, die das transformierte Plasmid durch homologe Rekombination am lysC-Lokus integriert haben. Nach Wachstum solcher Kolonien über Nacht in Medien, die kein Antibiotikum enthielten, wurden die Zellen auf ein Saccharose-CM-Agarmedium (10%

Saccharose) ausplattiert und bei 30°C für 24 Stunden inkubiert.

Da das im Vektor pCIS lysC thr311ile enthaltende sacB Gen Saccharose in ein toxisches Produkt umwandelt, können nur solche Kolonien anwachsen, die das sacB Gen durch einen zweiten homologen Rekombinationsschritt zwischen dem Wildtyp lysC Gen und dem mutierten Gen lysC thr311ile deletiert haben. Während der homologen Rekombination kann entweder das Wildtyp Gen oder das mutierte Gen zusammen mit dem sacB Gen deletiert werden. Wenn das sacB Gen zusammen mit dem Wildtyp Gen entfernt wird, resultiert eine mutierte Transformante.

Anwachsende Kolonien wurden gepickt, und auf einen Kanamycin-sensitiven Phänotyp hin untersucht. Klone mit deletiertem SacB Gen müssen gleichzeitig Kanamycin-sensitives Wachstumsverhalten zeigen. Solche Kan-sensitiven Klone wurden in einem Schüttelkolben auf ihre Lysin-Produktivität hin untersucht (siehe Beispiel 6). Zum Vergleich wurde der nichtbehandelte Stamm LU1479 angezogen. Klone mit einer gegenüber der Kontrolle erhöhten Lysin-Produktion wurden selektiert, chromosomale DNA wurde gewonnen und der entsprechende Bereich des lysC Gens wurde durch eine PCR-Reaktion amplifiziert und sequenziert. Ein solcher Klon mit der Eigenschaft erhöhter Lysin-Synthese und nachgewiesener Mutation in lysC an der Stelle 932 wurde mit LU1479 lysC 311ile bezeichnet).

Beispiel 5: Herstellung Ethionin-resistenter *C. glutamicum* Stämme

Im zweiten Schritt der Stammkonstruktion wurde der erhaltene Stamm LU1479 lysC 311ile (Beispiel 4) behandelt, um eine Ethionin-Resistenz (Kase, H. Nakayama K.Agr. Biol. Chem. 39 153-106 1975 L-methionine production by methionine analog-resistant mutants of *Corynebacterium glutamicum*) zu induzieren: Eine Übernachtskultur in BHI-Medium (Difco) wurde in Citratpuffer (50mM pH 5,5) gewaschen und bei 30°C für 20 min mit N-Methyl-nitrosoguanidin (10mg/ml in 50mM Citrat pH5,5) behandelt. Nach der Behandlung mit dem chemischen Mutagen N-Methyl-nitrosoguanidin wurden die Zellen gewaschen (Citratpuffer 50mM pH 5,5) und auf ein Medium plattiert, das aus folgenden Komponenten, berechnet auf 500ml, zusammengesetzt war: 10g $(\text{NH}_4)_2\text{SO}_4$, 0.5g KH_2PO_4 , 0.5g K_2HPO_4 , 0.125g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 21g MOPS, 50mg CaCl_2 , 15mg Proteokatechuat, 0,5mg Biotin, 1mg Thiamin, 5g/l D,L-Ethionin (Sigma Chemicals Deutschland), pH 7,0. Außerdem enthielt das Medium 0.5ml einer Spurensalzlösung aus: 10g/l $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 1g/l $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.1g/l $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 0.02g/l CuSO_4 , 0.002g/l $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$. Alle Salze wurden in 0,1M HCl gelöst. Das fertig zusammengestellte Medium wurde sterilfiltriert und nach Zugabe von 40ml steriler 50% Glucoselösung, mit flüssigem sterilem Agar in einer Endkonzentration von 1,5% Agar versetzt und in Kulturschalen ausgegossen.

Auf Platten mit dem beschriebenen Medium wurden mutagenisierte Zellen aufgebracht und 3-7 Tage bei 30°C inkubiert. Erhaltene Klone wurden isoliert, mindestens einmal auf dem Selektionsmedium vereinzelt und dann auf ihre Methionin-Produktivität in einem Schüttelkolben in Medium II untersucht (siehe Beispiel 6

5

Beispiel 6: Herstellung von Methionin mit dem Stamm LU1479 lysC 311ile ET-16.

Die in Beispiel 5 hergestellten Stämme wurden auf einer Agar-Platte mit CM-Medium für 2 Tag bei 30°C angezogen.

10 **CM-Agar:**

10,0 g/l D-Glucose, 2,5 g/l NaCl, 2,0 g/l Hamstoff, 10,0 g/l Bacto Pepton (Difco), 5,0 g/l Yeast Extract (Difco), 5,0 g/l Beef Extract (Difco), 22,0 g/l Agar (Difco), autoklaviert (20 min., 121°C)

15 Anschließend wurden die Zellen von der Platte abgekratzt und in Saline resuspendiert. Für die Hauptkultur wurden 10 ml Medium II und 0,5 g autoklaviertes CaCO₃ (Riedel de Haen) in einem 100 ml Erlenmeyerkolben mit der Zellsuspension bis zu einer OD_{600nm} von 1,5 beimpft und für 72h auf einem Orbitalschüttler mit 200 Upm bei 30°C inkubiert.

20 **Medium II:**

40g/l Saccharose

60g/l Melasse (auf 100% Zuckergehalt berechnet)

10g/l (NH₄)₂SO₄

0.4g/l MgSO₄*7H₂O

25 0.6g/l KH₂PO₄

0.3mg/l Thiamin*HCl

1mg/l Biotin (aus einer 1 mg/ml steril filtrierten Stammlösung die mit NH₄OH auf pH 8,0 eingestellt wurde)

2mg/l FeSO₄

30 2mg/l MnSO₄

mit NH₄OH auf pH 7,8 eingestellt, autoklaviert (121°C, 20 min). Zusätzlich wird Vitamin B12 (Hydroxycobalamin Sigma Chemicals) aus einer Stammlösung (200 µg/ml, steril filtriert) bis zu einer Endkonzentration von 100 µg/l zugegeben

35 Gebildetes Methionin, sowie andere Aminosäuren in der Kulturbrühe wurde mit Hilfe der Aminosäuresäure-Bestimmungsmethode von Agilent auf einer Agilent 1100 Series LC System

HPLC. Eine Derivatisierung vor der Säulentrennung mit Ortho-Phthalaldehyd erlaubte die Quantifizierung der gebildeten Aminosäuren. Die Auftrennung des Aminosäuregemisch fand auf einer Hypersil AA-Säule (Agilent) statt.

- 5 Solche Klone wurden isoliert, deren Methionin-Produktivität mindestens doppelt so hoch war, wie die des Ausgangsstamm LU1479 lysC 311ile. Ein solcher Klon wurde für die weiteren Versuche eingesetzt und bekam die Bezeichnung LU1479 lysC 311ile ET-16.

- 10 **Beispiel 7:** Klonierung von metY aus *Mycobacterium tuberculosis* und Klonierung in das Plasmid pC Phsdh metY_Mt

- 15 Chromosomale DNA von *Mycobacterium tuberculosis* wurde von der American Type Strain Culture Collection (ATCC, Atlanta-USA) aus dem Stamm ATCC 25584 bezogen. Chromosomale DNA aus *C. glutamicum* ATCC 13032 wurde nach Tauch et al. (1995) Plasmid 33:168-179 oder Eikmanns et al. (1994) Microbiology 140:1817-1828 präpariert.

- 20 Mit den Oligonukleotidprimer SEQ ID NO:75 und SEQ ID NO:76, der chromosomalen DNA aus *C. glutamicum* als Template und Pfu Turbo Polymerase (Fa. Stratagene) wurde mit Hilfe der Polymerase-Kettenreaktion (PCR) nach Standardmethoden wie Innis et al. (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press ein DNA Fragment von ca. 180 Basenpaaren aus dem nichtkodierenden 5'-Bereich (Promotorregion) der Homoserindehydrogenase (HsDH) amplifiziert. Das amplifizierte Fragment ist an seinem 5'-Ende von einer BamHI-Restriktionsschnittstelle und am 3'-Ende von einem über das Oligo eingeführten zu metY aus *Mycobacterium tuberculosis* homologen Bereich flankiert.

- 25 SEQ ID NO:75
5'-GAGAGGATCCGGAAGGTGAATCGAATTTCGG -3'
und
SEQ ID NO:76
30 5'-CTATTGCTGTCGGCGCTCATGATTCTCCAAAATAATCGC -3'

Das erhaltene DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt:

- 35 Ausgehend von der chromosomalen DNA aus *Mycobacterium tuberculosis* als Template für eine PCR Reaktion wurde mit den Oligonukleotidprimern SEQ ID NO:77 und SEQ ID NO:78 metY mit

Hilfe des GC-RICH PCR Systems (Roche Diagnostics, Mannheim) nach Angaben des Herstellers amplifiziert. Das amplifizierte Fragment ist an seinem 3'-Ende von einer XbaI-Restriktionsschnittstelle, die über das Oligo eingeführt wurde, flankiert.

5 SEQ ID NO:77

5'-ATGAGCGCCGACAGCAATAG -3'

und

SEQ ID NO:78

5'-GAACTCTAGATCAGAACGCCGCCACGGAC -3'

10

Das ca. 1,4 kb große erhaltene DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

15 In einer weiteren PCR Reaktion wurden die beiden oben erhaltenen Fragmente gemeinsam als Template eingesetzt. Durch die mit dem Oligonukleotidprimer SEQ ID NO:76 eingebrachte, zu dem metY Fragment homologen Bereichen, kommt es im Zuge der PCR-Reaktion zu einer Anlagerung beider Fragmente aneinander und einer Verlängerung zu einem durchgehenden DNA-Strang durch die eingesetzte Polymerase. Die Standardmethode wurde dahingehend modifiziert, dass die verwendeten Oligonukleotidprimer SEQ ID NO:75 und SEQ ID NO:78 erst mit
20 Beginn des 2. Zykluses dem Reaktionsansatz zugegeben wurden.

Das amplifizierte DNA Fragment von ungefähr 1,6 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit nach Angaben des Herstellers gereinigt. Im Anschluß daran wurde es mit den Restriktionsenzymen BamHI und XbaI (Roche Diagnostics, Mannheim) gespalten und ge-
25 lelektrophoretisch aufgetrennt. Anschließend wurde das ca. 1,6 kb große DNA Fragment mit GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aus der Agrose aufgereinigt.

Der Vektor pClik5MCS SEQ ID NO:65, im folgenden pC genannt, wurde mit den Restriktionsenzymen BamHI und XbaI (Roche Diagnostics, Mannheim) geschnitten und ein 5 kb großes Fragment nach elektrophoretischer Auftrennung mit GFX™PCR, DNA and Gel Band Purification Kit isoliert.
30

Das Vektorfragment wurde zusammen mit dem geschnittenen und aufgereinigten PCR-Fragment mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al.
35

(Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

5

Die Präparation der Plasmid DNA wurde nach Methoden und mit Materialien der Fa. Quiagen durchgeführt. Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

10

Das entstandene Plasmid pC Phsdh metY_Mt (*Mycobacterium tuberculosis*) ist als SEQ ID NO:79 aufgeführt. Die entsprechende Plasmidkarte ist in Figur 3 dargestellt.

Die Sequenz SEQ ID NO:79 umfasst die folgenden wesentlichen Teilbereiche:

15

LOCUS pC\Phsdh\metY_Mt 6591 bp DNA circular 21-JUL-2003

FEATURES Location/Qualifiers

20

CDS 156..1505
/vntifkey="4"
/label=metY\aus\M\tuberculosis

CDS 1855..2646

/vntifkey="4"

/label=Kan\R

25

CDS 4927..6048

/vntifkey="4"

/label=Rep\Protein

CDS 3919..4593

/vntifkey="4"

30

/label=ORF\1

CDS complement(2913..3773)

/vntifkey="4"

/label=Ori-EC\pMB)

35

Beispiel 8: Transformation des Stammes LU1479 lysC 311ile ET-16 mit dem Plasmid pC Phsdh metY_Mt

Der Stamm LU1479 lysC 311ile ET-16 wurde mit dem Plasmid pC Phsdh metY_Mt nach der

5 beschriebenen Methode (Liebl, et al. (1989) FEMS Microbiology Letters 53:299-303) transformiert. Die Transformationsmischung wurde auf CM-Platten plattiert, die zusätzlich 20mg/l Kanamycin enthielten, um eine Selektion auf Plasmid-haltige Zellen zu erreichen. Erhaltene Kan-resistente Klone wurden gepickt und vereinzelt. Die Methionin-Produktivität der Klone wurde in einem Schüttelkolbenversuch (s. Beispiel 6) untersucht. Der Stamm LU1479 lysC 311ile ET-16 pC Phsdh metY_Mt produzierte im Vergleich zu LU1479 lysC 311ile ET-16 signifikant mehr Methionin.

Patentansprüche

1. Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen
 5 Feinchemikalie, welches folgende Schritte umfasst:
- a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie
 produzierenden coryneformen Bakterienkultur, wobei in den coryneformen
 Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird,
 welche für ein Protein mit O-Acetyl-Homoserin-Sulphydrolase (metY) –Aktivität
 10 kodiert;
 - b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den
 Zellen der Bakterien, und
 - c) Isolieren der schwefelhaltigen Feinchemikalie.
2. Verfahren nach Anspruch 1, wobei die schwefelhaltige Feinchemikalie L-Methionin
 15 umfasst.
3. Verfahren nach einem der vorhergehenden Ansprüche, wobei die heterologe metY-
 kodierende Nukleotidsequenz zur metY-kodierenden Sequenz aus *Corynebacterium*
 20 *glutamicum* ATCC 13032 eine Sequenzhomologie vom weniger als 100% aufweist.
4. Verfahren nach Anspruch 3, wobei die metY-kodierende Sequenz aus einem der
 folgenden Organismen abgeleitet ist:

<i>Corynebacterium diphtheriae</i>	ATCC 14779
<i>Mycobacterium tuberculosis</i> CDC1551	ATCC 25584
<i>Clostridium acetobutylicum</i>	ATCC 824
<i>Bacillus halodurans</i>	ATCC21591
<i>Bacillus stearothermophilus</i>	ATCC 12980
<i>Chlorobium tepidum</i>	ATCC 49652
<i>Synechococcus</i> sp.	ATCC27104
<i>Emericella nidulans</i>	ATCC 36104
<i>Bacteroides fragilis</i>	ATCC 25285
<i>Lactococcus lactis</i>	ATCC 7962
<i>Bordetella bronchiseptica</i>	ATCC 19395
<i>Pseudomonas aeruginosa</i>	ATCC 17933
<i>Nitrosomonas europaea</i>	ATCC 19718
<i>Sinorhizobium meliloti</i>	ATCC 4399
<i>Thermotoga maritima</i>	ATCC 43589
<i>Streptococcus mutans</i>	ATCC 25175
<i>Burkholderia cepacia</i>	ATCC 25416
<i>Deinococcus radiodurans</i>	ATCC 13939

<i>Rhodobacter capsulatus</i>	ATCC 11166
<i>Pasteurella multocida</i>	ATCC 6530
<i>Clostridium difficile</i>	ATCC 9689
<i>Campylobacter jejuni</i>	ATCC 33560
<i>Streptococcus pneumoniae</i>	ATCC 6308
<i>Saccharomyces cerevisiae</i>	ATCC 2704
<i>Kluyveromyces lactis</i>	ATCC 8585
<i>Candida albicans</i>	ATCC 10231
<i>Schizosaccharomyces pombe</i>	ATCC 24969

5. Verfahren nach einem der vorhergehenden Ansprüche, wobei die metY-kodierende Sequenz eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51 und 53 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit metY-Aktivität kodiert, umfasst.
6. Verfahren nach einem der vorhergehenden Ansprüche, wobei die metY-kodierende Sequenz für ein Protein mit metY-Aktivität kodiert, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52 und 54 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metY-Aktivität steht, umfasst.
7. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metY-Sequenz eine in coryneformen Bakterien replizierbare oder eine stabil in das Chromosom integrierte DNA oder eine RNA ist.
8. Verfahren gemäß Anspruch 7, wobei man
 - a) einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden metY-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
 - b) einen Stamm einsetzt, in dem die kodierende metY-Sequenz in das Chromosom des Bakteriums integriert wurde
9. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metY-Sequenz überexprimiert wird.
10. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der

gewünschten schwefelhaltigen Feinchemikalie verstärkt ist oder derart mutiert ist, dass es durch Stoffwechselmetabolite nicht in seiner Aktivität beeinflusst wird.

5 11. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet ist, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.

12. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- 10
- a) dem für eine Aspartatkinase kodierenden Gen *lysC*,
 - b) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen *gap*,
 - c) dem für die 3-Phosphoglycerat Kinase kodierenden Gen *pgk*,
 - d) dem für die Pyruvat Carboxylase kodierenden Gen *pyc*,
 - 15 e) dem für die Triosephosphat Isomerase kodierenden Gen *tpi*,
 - f) dem für die Homoserin O-Acetyltransferase kodierenden Gen *metA*,
 - g) dem für die Cystathionin-gamma-Synthase kodierenden Gen *metB*,
 - h) dem für die Cystathionin-gamma-Lyase kodierenden Gen *metC*,
 - i) dem für die Serin-Hydroxymethyltransferase kodierenden Gen *glyA*,
 - 20 j) dem für die Methylen-Tetrahydrofolat-Reduktase kodierenden Gen *metF*,
 - k) dem für die Vitamin B12 abhängige Methionin-Synthase kodierenden Gen *methI*,
 - l) dem für die Phosphoserin-Aminotransferase kodierenden Gen *serC*
 - m) dem für die Phosphoserin-Phosphatase kodierenden Gen *serB*
 - n) dem für die Serine Acetyl-Transferase kodierenden Gen *cysE*, und
 - 25 o) dem für eine Homoserin-Dehydrogenase kodierenden Gen *hom*,

überexprimiert oder so mutiert ist, dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden.

30 13. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneformen Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- a) dem für die Homoserine-Kinase kodierenden Gen *thrB*,
- b) dem für die Threonin Dehydratase kodierenden Gen *ilvA*,
- 35 c) dem für die Threonin Synthase kodierenden Gen *thrC*
- d) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen *ddh*

- 5
- e) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen *pck*,
 - f) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen *pgi*,
 - g) dem für die Pyruvat-Oxidase kodierenden Gen *poxB*,
 - h) dem für die Dihydrodipicolinat Synthase kodierenden Gen *dapA*,
 - i) dem für die Dihydrodipicolinat Reduktase kodierenden Gen *dapB*; oder
 - j) dem für die Diaminopicolinat Decarboxylase kodierenden Gen

durch Veränderung der Expressionsrate oder durch Einführung einer gezielten Mutation abgeschwächt ist.

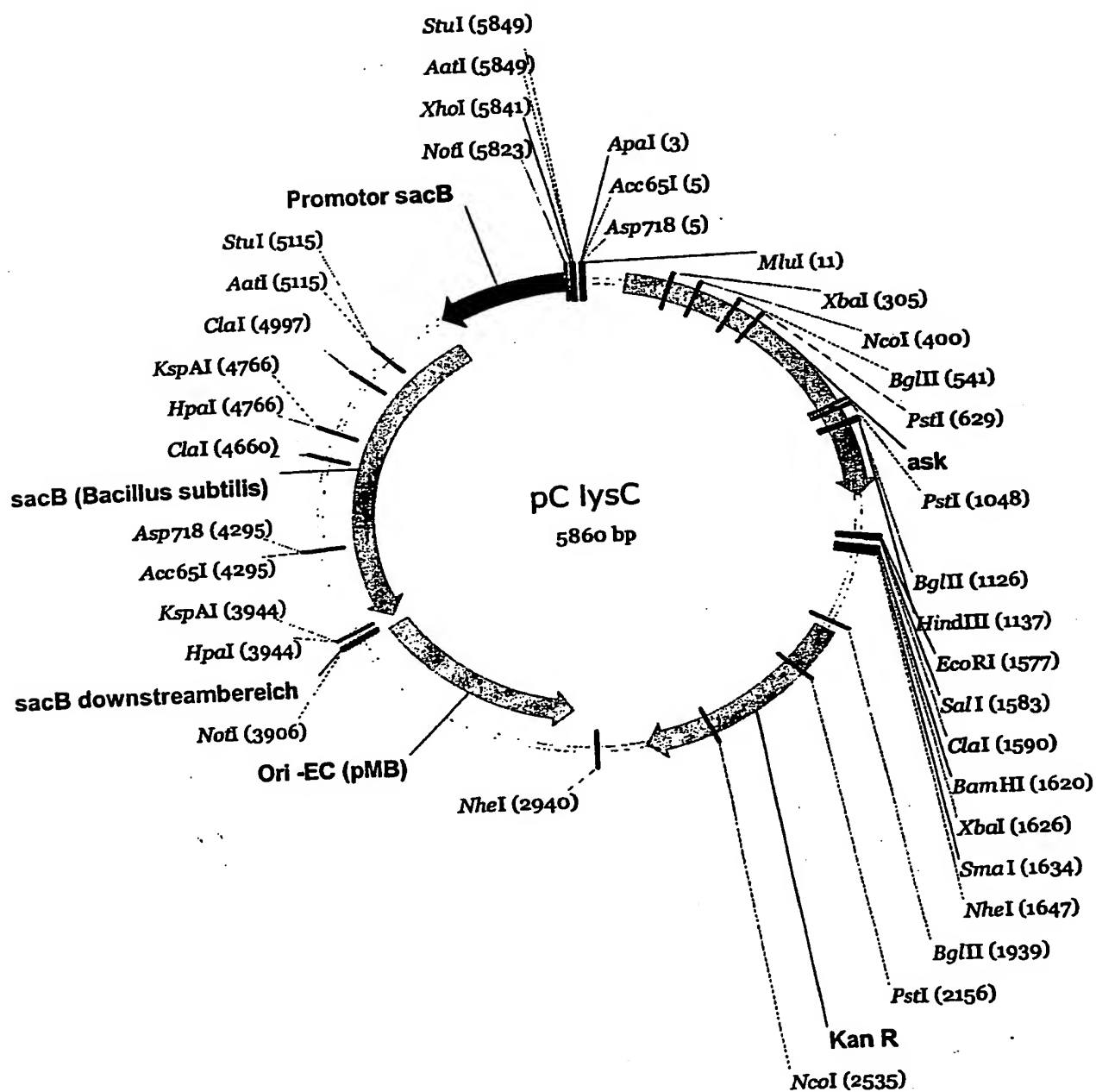
10 14. Verfahren gemäß einem oder mehreren der vorhergehenden Ansprüche, wobei man Mikroorganismen der Art *Corynebacterium glutamicum* einsetzt.

15 15. Verfahren zur Herstellung eines L-Methionin haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst

- 20
- a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
 - b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
 - c) Entfernung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
 - d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.

25 16. Verfahren gemäß Anspruch 15, wobei man Mikroorganismen gemäß der Definition in einem der Ansprüche 1 bis 14 einsetzt.

Fig. 1





2/3

Fig. 2

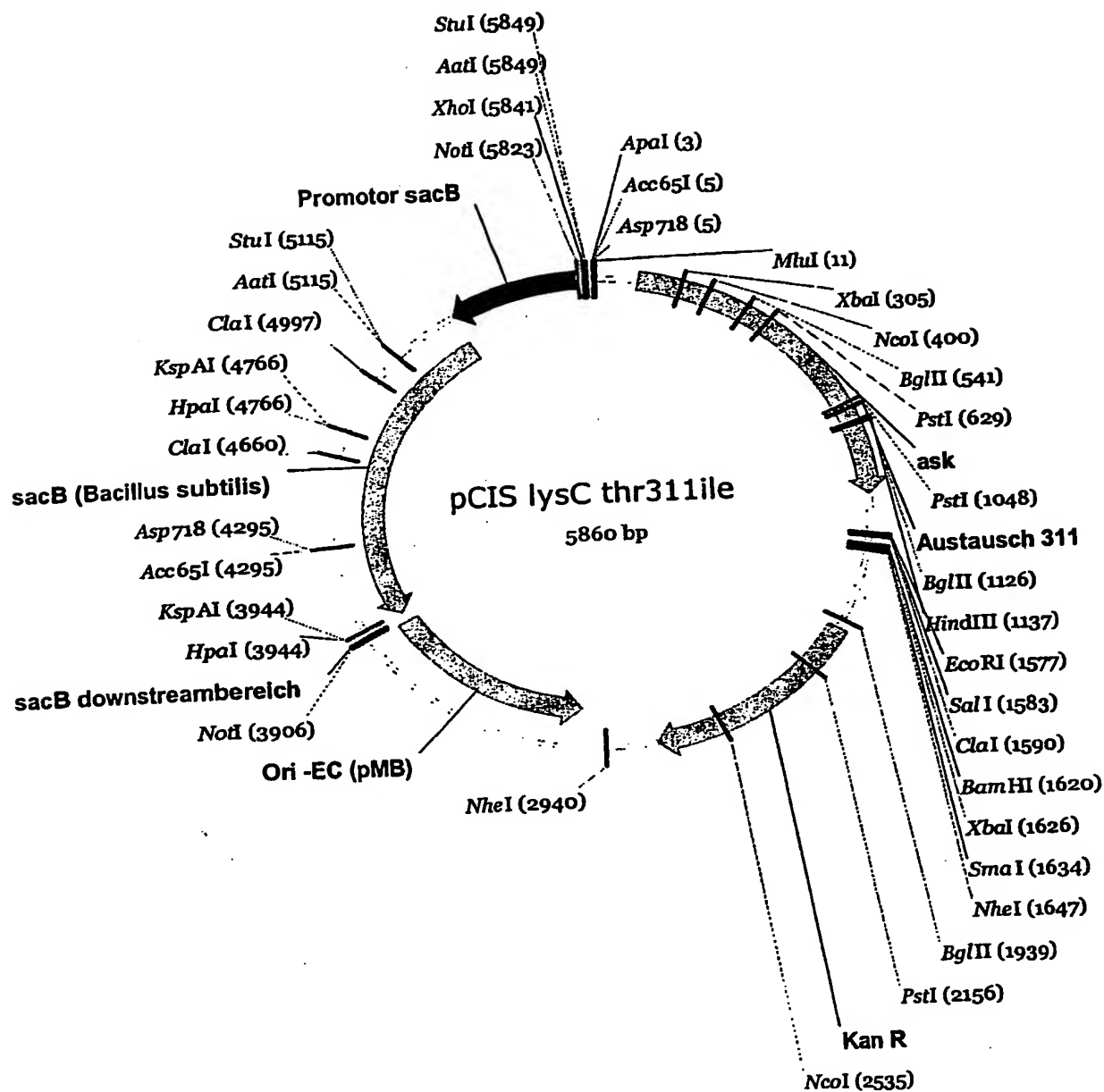
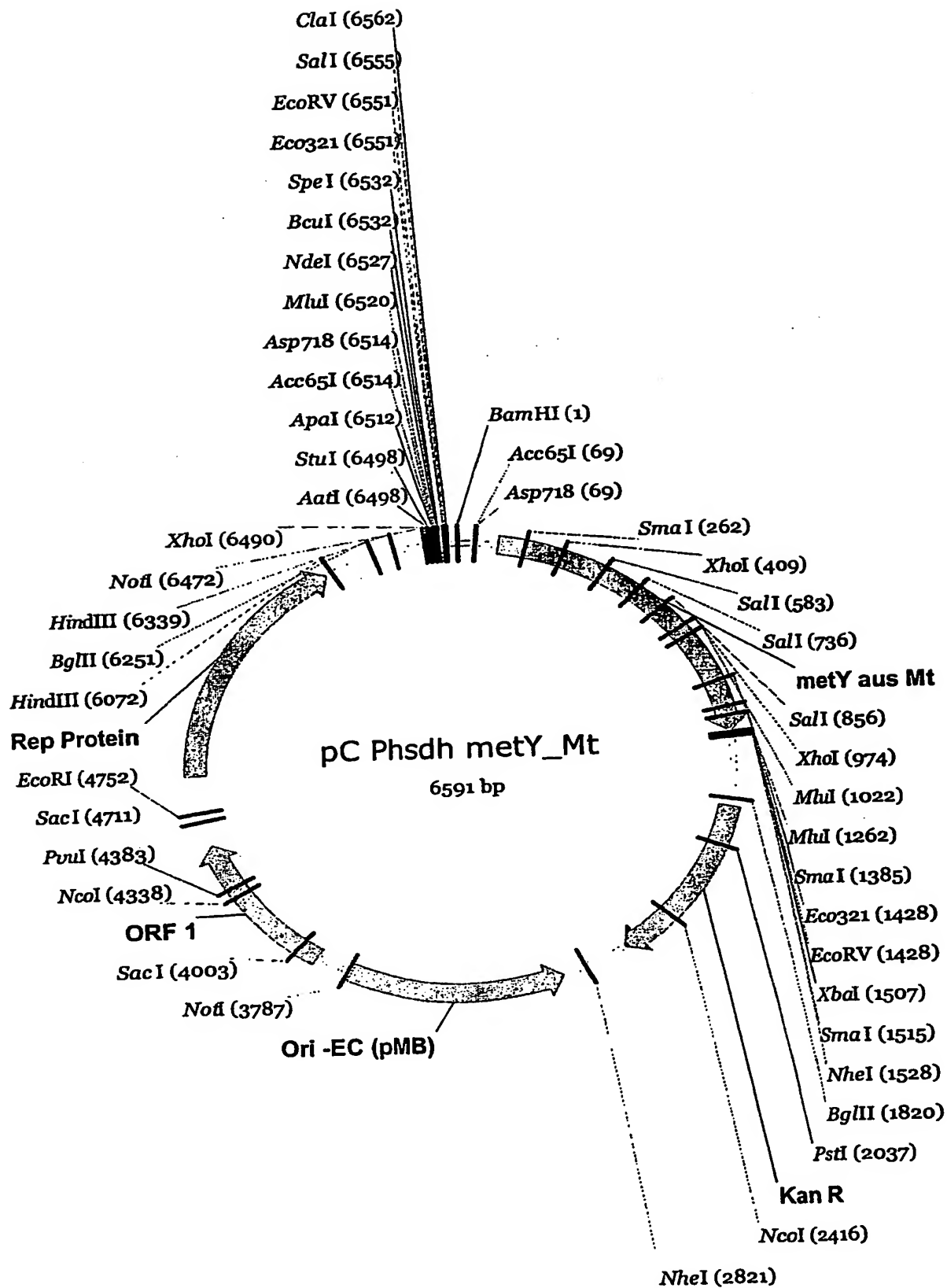




Fig. 3





SEQUENZPROTOKOLL

<110> BASF AG

<120> Verfahren zur Herstellung schwefelhaltiger Feinchemikalien

<130> M/43128

<160> 66

<170> PatentIn version 3.1

<210> 1

<211> 1317

<212> DNA

<213> Corynebacterium diphtheriae

<220>

<221> CDS

<222> (1)..(1317)

<223>

<400> 1

atg cca aca aaa tac gat aat tcc aat gcc aac aaa tgg ggt ttc gag	48
Met Pro Thr Lys Tyr Asp Asn Ser Asn Ala Asn Lys Trp Gly Phe Glu	
1 5 10 15	
act cgc tcc atc cac gca gga caa agc gtc gat agt gat acc agt gcc	96
Thr Arg Ser Ile His Ala Gly Gln Ser Val Asp Ser Asp Thr Ser Ala	
20 25 30	
cgc aac cta ccg att tac ctg aca tca tcg tac gtt ttt aat gac gcc	144
Arg Asn Leu Pro Ile Tyr Leu Thr Ser Ser Tyr Val Phe Asn Asp Ala	
35 40 45	
gaa cac gca gca aac cgc ttc aac ctt tcc gac gcc ggc ccg gtt tac	192
Glu His Ala Ala Asn Arg Phe Asn Leu Ser Asp Ala Gly Pro Val Tyr	
50 55 60	
tct cgc ctg acc aac cca act gtc gcg gca gtc gaa gaa cgc cta gcc	240
Ser Arg Leu Thr Asn Pro Thr Val Ala Ala Val Glu Glu Arg Leu Ala	
65 70 75 80	

aat ctt gaa ggt ggc gta cac gcc gta ctt ttc gct tcc gga atg gcc	288
Asn Leu Glu Gly Gly Val His Ala Val Leu Phe Ala Ser Gly Met Ala	
85 90 95	
gcc gaa acc gcc gca atc ctc aac atc gcc cgc gcg ggt tcc cac atc	336
Ala Glu Thr Ala Ala Ile Leu Asn Ile Ala Arg Ala Gly Ser His Ile	
100 105 110	
gtg tcc agt cct cgc att tac ggc ggc acc gaa aca ctc ttt gcc gtc	384
Val Ser Ser Pro Arg Ile Tyr Gly Gly Thr Glu Thr Leu Phe Ala Val	
115 120 125	
aca ttg gca cgc ctg ggc atc gaa acc act ttc gta gaa aat cct gac	432
Thr Leu Ala Arg Leu Gly Ile Glu Thr Thr Phe Val Glu Asn Pro Asp	
130 135 140	
gac cca gcc tca tgg gag gct gca gtt caa gac aac acg gta gct ctc	480
Asp Pro Ala Ser Trp Glu Ala Ala Val Gln Asp Asn Thr Val Ala Leu	
145 150 155 160	
tac gga gaa acc ttc gct aat cca caa gca gac gtg ctt gat att ccc	528
Tyr Gly Glu Thr Phe Ala Asn Pro Gln Ala Asp Val Leu Asp Ile Pro	
165 170 175	
gca atc gca gag gtt gcc cat aaa cat caa gta cca ctg atc gtc gac	576
Ala Ile Ala Glu Val Ala His Lys His Gln Val Pro Leu Ile Val Asp	
180 185 190	
aac acc ctc gca acc gca gcc ctt gta cgc ccc ctc gaa ctc ggt gca	624
Asn Thr Leu Ala Thr Ala Ala Leu Val Arg Pro Leu Glu Leu Gly Ala	
195 200 205	
gac gtc gtc gtg gca tcc cta acc aag ttc tac acc gga aat ggc tcc	672
Asp Val Val Val Ala Ser Leu Thr Lys Phe Tyr Thr Gly Asn Gly Ser	
210 215 220	
gga ctc ggc gga gtg ctt atc gac ggc gga aac ttc gac tgg acc gtc	720
Gly Leu Gly Gly Val Leu Ile Asp Gly Gly Asn Phe Asp Trp Thr Val	
225 230 235 240	
aca cgc aac ggc gaa ccg atc ttc ccc gac ttt gtc acc cca gat ccc	768
Thr Arg Asn Gly Glu Pro Ile Phe Pro Asp Phe Val Thr Pro Asp Pro	
245 250 255	
gcc tat cac ggt ctc aag tat tcc gat ctt ggt gcc ccc gcc ttc gga	816
Ala Tyr His Gly Leu Lys Tyr Ser Asp Leu Gly Ala Pro Ala Phe Gly	
260 265 270	
cta aag gct cgc gtc gga ctc ctg cgc gac acc ggc gca gcc cca tca	864
Leu Lys Ala Arg Val Gly Leu Leu Arg Asp Thr Gly Ala Ala Pro Ser	
275 280 285	
cca ctc aac gca tgg atc acc gca caa ggg ctc gac acc ctc tcg cta	912
Pro Leu Asn Ala Trp Ile Thr Ala Gln Gly Leu Asp Thr Leu Ser Leu	
290 295 300	
cga gta caa cgc cac aac gaa aac gca ctc gca gta gca caa ttc ctc	960
Arg Val Gln Arg His Asn Glu Asn Ala Leu Ala Val Ala Gln Phe Leu	
305 310 315 320	
gcc aac cac gag aaa gta gcc aag gtt aac tac gca ggc ctt ccc gac	1008
Ala Asn His Glu Lys Val Ala Lys Val Asn Tyr Ala Gly Leu Pro Asp	

325	330	335	
tcc cct tgg tac cca gtc aaa gaa aaa ctc gga ttc gac tac acc ggc Ser Pro Trp Tyr Pro Val Lys Glu Lys Leu Gly Phe Asp Tyr Thr Gly 340 345 350			1056
tcc gta ctt tcc ttt gac gtt aaa ggt gga aaa aac gaa gca tgg cgc Ser Val Leu Ser Phe Asp Val Lys Gly Gly Lys Asn Glu Ala Trp Arg 355 360 365			1104
ttt atc gac gca ctc aaa cta cac tcg aac ctc gcc aac gtc gga gac Phe Ile Asp Ala Leu Lys Leu His Ser Asn Leu Ala Asn Val Gly Asp 370 375 380			1152
gta cgt tcc ctc gta gtc cac cca gcg acc acc acg cac tca caa tcg Val Arg Ser Leu Val Val His Pro Ala Thr Thr Thr His Ser Gln Ser 385 390 395 400			1200
gaa gaa tcg gca ctt cta gcc gca gga att aat caa gca acc atc cga Glu Glu Ser Ala Leu Leu Ala Ala Gly Ile Asn Gln Ala Thr Ile Arg 405 410 415			1248
ctc tcc gtc ggc atc gaa tcc atc gac gac atc atc gcc gac ctc aca Leu Ser Val Gly Ile Glu Ser Ile Asp Asp Ile Ile Ala Asp Leu Thr 420 425 430			1296
gca ggt ttc gac gca atc taa Ala Gly Phe Asp Ala Ile 435			1317

<210> 2

<211> 438

<212> PRT

<213> Corynebacterium diphtheriae

<400> 2

Met	Pro	Thr	Lys	Tyr	Asp	Asn	Ser	Asn	Ala	Asn	Lys	Trp	Gly	Phe	Glu
1				5					10					15	

Thr	Arg	Ser	Ile	His	Ala	Gly	Gln	Ser	Val	Asp	Ser	Asp	Thr	Ser	Ala
			20					25					30		

Arg	Asn	Leu	Pro	Ile	Tyr	Leu	Thr	Ser	Ser	Tyr	Val	Phe	Asn	Asp	Ala
		35					40					45			

Glu	His	Ala	Ala	Asn	Arg	Phe	Asn	Leu	Ser	Asp	Ala	Gly	Pro	Val	Tyr
	50					55					60				

Ser	Arg	Leu	Thr	Asn	Pro	Thr	Val	Ala	Ala	Val	Glu	Glu	Arg	Leu	Ala
65					70					75					80

Asn Leu Glu Gly Gly Val His Ala Val Leu Phe Ala Ser Gly Met Ala
85 90 95

Ala Glu Thr Ala Ala Ile Leu Asn Ile Ala Arg Ala Gly Ser His Ile
100 105 110

Val Ser Ser Pro Arg Ile Tyr Gly Gly Thr Glu Thr Leu Phe Ala Val
115 120 125

Thr Leu Ala Arg Leu Gly Ile Glu Thr Thr Phe Val Glu Asn Pro Asp
130 135 140

Asp Pro Ala Ser Trp Glu Ala Ala Val Gln Asp Asn Thr Val Ala Leu
145 150 155 160

Tyr Gly Glu Thr Phe Ala Asn Pro Gln Ala Asp Val Leu Asp Ile Pro
165 170 175

Ala Ile Ala Glu Val Ala His Lys His Gln Val Pro Leu Ile Val Asp
180 185 190

Asn Thr Leu Ala Thr Ala Ala Leu Val Arg Pro Leu Glu Leu Gly Ala
195 200 205

Asp Val Val Val Ala Ser Leu Thr Lys Phe Tyr Thr Gly Asn Gly Ser
210 215 220

Gly Leu Gly Gly Val Leu Ile Asp Gly Gly Asn Phe Asp Trp Thr Val
225 230 235 240

Thr Arg Asn Gly Glu Pro Ile Phe Pro Asp Phe Val Thr Pro Asp Pro
245 250 255

Ala Tyr His Gly Leu Lys Tyr Ser Asp Leu Gly Ala Pro Ala Phe Gly
260 265 270

Leu Lys Ala Arg Val Gly Leu Leu Arg Asp Thr Gly Ala Ala Pro Ser
275 280 285

Pro Leu Asn Ala Trp Ile Thr Ala Gln Gly Leu Asp Thr Leu Ser Leu
290 295 300

Arg Val Gln Arg His Asn Glu Asn Ala Leu Ala Val Ala Gln Phe Leu
305 310 315 320

Ala Asn His Glu Lys Val Ala Lys Val Asn Tyr Ala Gly Leu Pro Asp
325 330 335

Ser Pro Trp Tyr Pro Val Lys Glu Lys Leu Gly Phe Asp Tyr Thr Gly
 340 345 350

Ser Val Leu Ser Phe Asp Val Lys Gly Gly Lys Asn Glu Ala Trp Arg
 355 360 365

Phe Ile Asp Ala Leu Lys Leu His Ser Asn Leu Ala Asn Val Gly Asp
 370 375 380

Val Arg Ser Leu Val Val His Pro Ala Thr Thr Thr His Ser Gln Ser
 385 390 395 400

Glu Glu Ser Ala Leu Leu Ala Ala Gly Ile Asn Gln Ala Thr Ile Arg
 405 410 415

Leu Ser Val Gly Ile Glu Ser Ile Asp Asp Ile Ile Ala Asp Leu Thr
 420 425 430

Ala Gly Phe Asp Ala Ile
 435

<210> 3

<211> 1350

<212> DNA

<213> Mycobacterium tuberculosis

<220>

<221> CDS

<222> (1) .. (1350)

<223>

<400> 3
 atg agc gcc gac agc aat agc acc gac gcc gat ccg acc gcg cat tgg 48
 Met Ser Ala Asp Ser Asn Ser Thr Asp Ala Asp Pro Thr Ala His Trp
 1 5 10 15

tcg ttc gaa acc aaa cag ata cac gct ggt cag cac cct gat ccg acc 96
 Ser Phe Glu Thr Lys Gln Ile His Ala Gly Gln His Pro Asp Pro Thr
 20 25 30

acc aac gcc cgg gct ctg ccg atc tat gcg acc acg tcg tac acc ttc 144
 Thr Asn Ala Arg Ala Leu Pro Ile Tyr Ala Thr Thr Ser Tyr Thr Phe
 35 40 45

gac gac acc gcg cac gcc gcc gcc ctg ttc gga ctg gaa att ccg ggc Asp Asp Thr Ala His Ala Ala Ala Leu Phe Gly Leu Glu Ile Pro Gly 50 55 60	192
aat atc tac acc cgg atc ggc aac ccc acc acc gac gtc gtc gag cag Asn Ile Tyr Thr Arg Ile Gly Asn Pro Thr Thr Asp Val Val Glu Gln 65 70 75 80	240
cgc atc gcc gcg ctc gag ggc ggt gtg gcc gcg ctg ttc ctg tcg tcg Arg Ile Ala Ala Leu Glu Gly Gly Val Ala Ala Leu Phe Leu Ser Ser 85 90 95	288
ggg cag gcc gcg gag acg ttc gcc atc ttg aac ctg gcc ggc gcg ggc Gly Gln Ala Ala Glu Thr Phe Ala Ile Leu Asn Leu Ala Gly Ala Gly 100 105 110	336
gat cac atc gtg tcc agc ccg cgc ctg tac ggc ggc acc tac aac ctg Asp His Ile Val Ser Ser Pro Arg Leu Tyr Gly Gly Thr Tyr Asn Leu 115 120 125	384
ttc cac tat tcg ctg gcc aag ctc ggc atc gag gtc agc ttc gtc gac Phe His Tyr Ser Leu Ala Lys Leu Gly Ile Glu Val Ser Phe Val Asp 130 135 140	432
gat ccg gac gat ctg gac acc tgg cag gcg gcg gta cgg ccc aac acc Asp Pro Asp Asp Leu Asp Thr Trp Gln Ala Ala Val Arg Pro Asn Thr 145 150 155 160	480
aag gcg ttc ttc gcc gag acc atc tcc aac ccg cag atc gac ctg ctg Lys Ala Phe Phe Ala Glu Thr Ile Ser Asn Pro Gln Ile Asp Leu Leu 165 170 175	528
gac acc ccg gcg gtt tcc gag gtc gcc cat cgc aac ggg gtg ccg ttg Asp Thr Pro Ala Val Ser Glu Val Ala His Arg Asn Gly Val Pro Leu 180 185 190	576
atc gtc gac aac acc atc gcc acg cca tac ctg atc caa ccg ttg gcc Ile Val Asp Asn Thr Ile Ala Thr Pro Tyr Leu Ile Gln Pro Leu Ala 195 200 205	624
cag ggc gcc gac atc gtc gtg cat tcg gcc acc aag tac ctg ggc ggc Gln Gly Ala Asp Ile Val Val His Ser Ala Thr Lys Tyr Leu Gly Gly 210 215 220	672
cac ggt gcc gcc atc gcg ggt gtg atc gtc gac ggc ggc aac ttc gat His Gly Ala Ala Ile Ala Gly Val Ile Val Asp Gly Gly Asn Phe Asp 225 230 235 240	720
tgg acc cag ggc cgc ttc ccc ggc ttc acc acc ccc gac ccc agc tac Trp Thr Gln Gly Arg Phe Pro Gly Phe Thr Thr Pro Asp Pro Ser Tyr 245 250 255	768
cac ggc gtg gtg ttc gcc gag ctg ggt cca ccg gcg ttt gcg ctc aaa His Gly Val Val Phe Ala Glu Leu Gly Pro Pro Ala Phe Ala Leu Lys 260 265 270	816
gct cga gtg cag ctg ctc cgt gac tac ggc tcg gcg gct tcg ccg ttc Ala Arg Val Gln Leu Leu Arg Asp Tyr Gly Ser Ala Ala Ser Pro Phe 275 280 285	864
aac gcg ttc ttg gtg gcg cag ggt ctg gaa acg ctg agc ctg cgg atc Asn Ala Phe Leu Val Ala Gln Gly Leu Glu Thr Leu Ser Leu Arg Ile 290 295 300	912

gag cgg cac gtc gcc aac gcg cag cgc gtc gcc gag ttc ctg gcc gcc	960
Glu Arg His Val Ala Asn Ala Gln Arg Val Ala Glu Phe Leu Ala Ala	
305 310 315 320	
cgc gac gac gtg ctt tcg gtc aac tat gcg ggg ctg ccc tcc tcg ccc	1008
Arg Asp Asp Val Leu Ser Val Asn Tyr Ala Gly Leu Pro Ser Ser Pro	
325 330 335	
tgg cat gag cgg gcc aag agg ctg gcg ccc aag gga acc ggg gcc gtg	1056
Trp His Glu Arg Ala Lys Arg Leu Ala Pro Lys Gly Thr Gly Ala Val	
340 345 350	
ctg tcc ttc gag ttg gcc ggc ggc atc gag gcc ggc aag gca ttc gtg	1104
Leu Ser Phe Glu Leu Ala Gly Gly Ile Glu Ala Gly Lys Ala Phe Val	
355 360 365	
aac gcg ttg aag ctg cac agc cac gtc gcc aac atc ggt gac gtg cgc	1152
Asn Ala Leu Lys Leu His Ser His Val Ala Asn Ile Gly Asp Val Arg	
370 375 380	
tcg ctg gtg atc cac ccg gca tcg acc act cat gcc cag ctg agc ccg	1200
Ser Leu Val Ile His Pro Ala Ser Thr Thr His Ala Gln Leu Ser Pro	
385 390 395 400	
gcc gag cag ctg gcg acc ggg gtc agc ccg ggc ctg gtg cgt ttg gct	1248
Ala Glu Gln Leu Ala Thr Gly Val Ser Pro Gly Leu Val Arg Leu Ala	
405 410 415	
gtg ggc atc gaa ggt atc gac gat atc ctg gcc gac ctg gag ctt ggc	1296
Val Gly Ile Glu Gly Ile Asp Asp Ile Leu Ala Asp Leu Glu Leu Gly	
420 425 430	
ttt gcc gcg gcc cgc aga ttc agc gcc gac ccg cag tcc gtg gcg gcg	1344
Phe Ala Ala Ala Arg Arg Phe Ser Ala Asp Pro Gln Ser Val Ala Ala	
435 440 445	
ttc tga	1350
Phe	

<210> 4

<211> 449

<212> PRT

<213> Mycobacterium tuberculosis

<400> 4

Met Ser Ala Asp Ser Asn Ser Thr Asp Ala Asp Pro Thr Ala His Trp
1 5 10 15

Ser Phe Glu Thr Lys Gln Ile His Ala Gly Gln His Pro Asp Pro Thr
20 25 30

Thr Asn Ala Arg Ala Leu Pro Ile Tyr Ala Thr Thr Ser Tyr Thr Phe

35

40

45

Asp Asp Thr Ala His Ala Ala Ala Leu Phe Gly Leu Glu Ile Pro Gly
50 55 60

Asn Ile Tyr Thr Arg Ile Gly Asn Pro Thr Thr Asp Val Val Glu Gln
65 70 75 80

Arg Ile Ala Ala Leu Glu Gly Gly Val Ala Ala Leu Phe Leu Ser Ser
85 90 95

Gly Gln Ala Ala Glu Thr Phe Ala Ile Leu Asn Leu Ala Gly Ala Gly
100 105 110

Asp His Ile Val Ser Ser Pro Arg Leu Tyr Gly Gly Thr Tyr Asn Leu
115 120 125

Phe His Tyr Ser Leu Ala Lys Leu Gly Ile Glu Val Ser Phe Val Asp
130 135 140

Asp Pro Asp Asp Leu Asp Thr Trp Gln Ala Ala Val Arg Pro Asn Thr
145 150 155 160

Lys Ala Phe Phe Ala Glu Thr Ile Ser Asn Pro Gln Ile Asp Leu Leu
165 170 175

Asp Thr Pro Ala Val Ser Glu Val Ala His Arg Asn Gly Val Pro Leu
180 185 190

Ile Val Asp Asn Thr Ile Ala Thr Pro Tyr Leu Ile Gln Pro Leu Ala
195 200 205

Gln Gly Ala Asp Ile Val Val His Ser Ala Thr Lys Tyr Leu Gly Gly
210 215 220

His Gly Ala Ala Ile Ala Gly Val Ile Val Asp Gly Gly Asn Phe Asp
225 230 235 240

Trp Thr Gln Gly Arg Phe Pro Gly Phe Thr Thr Pro Asp Pro Ser Tyr
245 250 255

His Gly Val Val Phe Ala Glu Leu Gly Pro Pro Ala Phe Ala Leu Lys
260 265 270

Ala Arg Val Gln Leu Leu Arg Asp Tyr Gly Ser Ala Ala Ser Pro Phe
275 280 285

Asn Ala Phe Leu Val Ala Gln Gly Leu Glu Thr Leu Ser Leu Arg Ile
290 295 300

Glu Arg His Val Ala Asn Ala Gln Arg Val Ala Glu Phe Leu Ala Ala
305 310 315 320

Arg Asp Asp Val Leu Ser Val Asn Tyr Ala Gly Leu Pro Ser Ser Pro
325 330 335

Trp His Glu Arg Ala Lys Arg Leu Ala Pro Lys Gly Thr Gly Ala Val
340 345 350

Leu Ser Phe Glu Leu Ala Gly Gly Ile Glu Ala Gly Lys Ala Phe Val
355 360 365

Asn Ala Leu Lys Leu His Ser His Val Ala Asn Ile Gly Asp Val Arg
370 375 380

Ser Leu Val Ile His Pro Ala Ser Thr Thr His Ala Gln Leu Ser Pro
385 390 395 400

Ala Glu Gln Leu Ala Thr Gly Val Ser Pro Gly Leu Val Arg Leu Ala
405 410 415

Val Gly Ile Glu Gly Ile Asp Asp Ile Leu Ala Asp Leu Glu Leu Gly
420 425 430

Phe Ala Ala Ala Arg Arg Phe Ser Ala Asp Pro Gln Ser Val Ala Ala
435 440 445

Phe

<210> 5

<211> 1284

<212> DNA

<213> Clostridium acetobutylicum

<220>

<221> CDS

<222> (1) .. (1284)

<223>

<400> 5
atg agt gaa gaa aga aaa ttt ggt ttt gaa aca tta cag gtt cat gca 48
Met Ser Glu Glu Arg Lys Phe Gly Phe Glu Thr Leu Gln Val His Ala
1 5 10 15

gga caa gtt gct gat cca act aca gga tca aga gct gta cct att tat 96
Gly Gln Val Ala Asp Pro Thr Thr Gly Ser Arg Ala Val Pro Ile Tyr
20 25 30

caa aca aca tca tat gta ttt aaa aat gct gat cat gca gca aat tta 144
Gln Thr Thr Ser Tyr Val Phe Lys Asn Ala Asp His Ala Ala Asn Leu
35 40 45

ttt caa ttg aaa gaa cct gga aat gta tat aca agg ata atg aat cca 192
Phe Gln Leu Lys Glu Pro Gly Asn Val Tyr Thr Arg Ile Met Asn Pro
50 55 60

aca act gat gta ttt gaa caa aga gta gca gct ctt gag ggc gga gtt 240
Thr Thr Asp Val Phe Glu Gln Arg Val Ala Ala Leu Glu Gly Gly Val
65 70 75 80

gct gga ctt gca aca gca tca gga ctt gca gca att acc tat gct att 288
Ala Gly Leu Ala Thr Ala Ser Gly Leu Ala Ala Ile Thr Tyr Ala Ile
85 90 95

tta aat gtg gca agt gct ggg gat gaa att gtt gca gca agt acc tta 336
Leu Asn Val Ala Ser Ala Gly Asp Glu Ile Val Ala Ala Ser Thr Leu
100 105 110

tat ggt gga aca tat gaa tta ttt ggg gtt act ctt aag aag ctt gga 384
Tyr Gly Gly Thr Tyr Glu Leu Phe Gly Val Thr Leu Lys Lys Leu Gly
115 120 125

ata aag gtt gtt ttt gta gat cca gat aat cct gaa aat ata aga aaa 432
Ile Lys Val Val Phe Val Asp Pro Asp Asn Pro Glu Asn Ile Arg Lys
130 135 140

gca ata aat gat agg aca aaa gct gta tat ggg gaa act att gga aat 480
Ala Ile Asn Asp Arg Thr Lys Ala Val Tyr Gly Glu Thr Ile Gly Asn
145 150 155 160

cca aga ata aat gtt ttg gat ata gag gca gta gct aaa att gcc cat 528
Pro Arg Ile Asn Val Leu Asp Ile Glu Ala Val Ala Lys Ile Ala His
165 170 175

gaa aat aaa ata cca ctt ata atc gat aat aca ttt ggt aca ccg tat 576
Glu Asn Lys Ile Pro Leu Ile Ile Asp Asn Thr Phe Gly Thr Pro Tyr
180 185 190

ctt ata aga cct ata gaa ttt gga gca gat ata gtt gta cat tca gca 624
Leu Ile Arg Pro Ile Glu Phe Gly Ala Asp Ile Val Val His Ser Ala
195 200 205

aca aag ttt ata gga gga cat gga act act ata ggt gga att ata gtt 672
Thr Lys Phe Ile Gly Gly His Gly Thr Thr Ile Gly Gly Ile Ile Val
210 215 220

gat ggt gga aaa ttt gat tgg aga gct agt gga aag ttt cct gat ttt 720
Asp Gly Gly Lys Phe Asp Trp Arg Ala Ser Gly Lys Phe Pro Asp Phe
225 230 235 240

aca aca ccg gat aag agc tat aat gga ctt ata tat gct gat cta ggt 768
 Thr Thr Pro Asp Lys Ser Tyr Asn Gly Leu Ile Tyr Ala Asp Leu Gly
 245 250 255

gca cct gct ttt gct tta aaa gca aga gtt caa ctt tta aga aat aca 816
 Ala Pro Ala Phe Ala Leu Lys Ala Arg Val Gln Leu Leu Arg Asn Thr
 260 265 270

ggg gca acg ctt agt cca caa agt gct ttt tat ttc cta caa ggg ttg 864
 Gly Ala Thr Leu Ser Pro Gln Ser Ala Phe Tyr Phe Leu Gln Gly Leu
 275 280 285

gaa tca ctt tca ctt agg gtt caa aaa cat gtt gat aat aca aga aag 912
 Glu Ser Leu Ser Leu Arg Val Gln Lys His Val Asp Asn Thr Arg Lys
 290 295 300

gta gtt gaa ttc ttg aag aac cat cca aaa gtt tca tgg ata aat tat 960
 Val Val Glu Phe Leu Lys Asn His Pro Lys Val Ser Trp Ile Asn Tyr
 305 310 315 320

cct gaa ctt gag gaa agt cct tat aaa gag tta gca aat aaa tat ctt 1008
 Pro Glu Leu Glu Glu Ser Pro Tyr Lys Glu Leu Ala Asn Lys Tyr Leu
 325 330 335

cca aag ggt gca ggc tca ata ttt aca ttt gga ata aag gga gga ctt 1056
 Pro Lys Gly Ala Gly Ser Ile Phe Thr Phe Gly Ile Lys Gly Gly Leu
 340 345 350

gaa gct ggt aaa aga ttt ata aat agt gtt aaa cta ttc tct ctt ttg 1104
 Glu Ala Gly Lys Arg Phe Ile Asn Ser Val Lys Leu Phe Ser Leu Leu
 355 360 365

gca aat gtt gca gat gca aaa tca ctt gtt ata cat cct tca agt aca 1152
 Ala Asn Val Ala Asp Ala Lys Ser Leu Val Ile His Pro Ser Ser Thr
 370 375 380

act cat gct gaa ctt aat gaa gaa gaa caa aaa gca gct ggt gtt act 1200
 Thr His Ala Glu Leu Asn Glu Glu Glu Gln Lys Ala Ala Gly Val Thr
 385 390 395 400

cca gat atg ata aga ctt tca ata gga gta gag gat gca gag gat tta 1248
 Pro Asp Met Ile Arg Leu Ser Ile Gly Val Glu Asp Ala Glu Asp Leu
 405 410 415

ata tgg gac tta aat caa gct ctc gaa caa gct taa 1284
 Ile Trp Asp Leu Asn Gln Ala Leu Glu Gln Ala
 420 425

<210> 6

<211> 427

<212> PRT

<213> Clostridium acetobutylicum

<400> 6

Met Ser Glu Glu Arg Lys Phe Gly Phe Glu Thr Leu Gln Val His Ala
 1 5 10 15

Gly Gln Val Ala Asp Pro Thr Thr Gly Ser Arg Ala Val Pro Ile Tyr
 20 25 30

Gln Thr Thr Ser Tyr Val Phe Lys Asn Ala Asp His Ala Ala Asn Leu
 35 40 45

Phe Gln Leu Lys Glu Pro Gly Asn Val Tyr Thr Arg Ile Met Asn Pro
 50 55 60

Thr Thr Asp Val Phe Glu Gln Arg Val Ala Ala Leu Glu Gly Gly Val
 65 70 75 80

Ala Gly Leu Ala Thr Ala Ser Gly Leu Ala Ala Ile Thr Tyr Ala Ile
 85 90 95

Leu Asn Val Ala Ser Ala Gly Asp Glu Ile Val Ala Ala Ser Thr Leu
 100 105 110

Tyr Gly Gly Thr Tyr Glu Leu Phe Gly Val Thr Leu Lys Lys Leu Gly
 115 120 125

Ile Lys Val Val Phe Val Asp Pro Asp Asn Pro Glu Asn Ile Arg Lys
 130 135 140

Ala Ile Asn Asp Arg Thr Lys Ala Val Tyr Gly Glu Thr Ile Gly Asn
 145 150 155 160

Pro Arg Ile Asn Val Leu Asp Ile Glu Ala Val Ala Lys Ile Ala His
 165 170 175

Glu Asn Lys Ile Pro Leu Ile Ile Asp Asn Thr Phe Gly Thr Pro Tyr
 180 185 190

Leu Ile Arg Pro Ile Glu Phe Gly Ala Asp Ile Val Val His Ser Ala
 195 200 205

Thr Lys Phe Ile Gly Gly His Gly Thr Thr Ile Gly Gly Ile Ile Val
 210 215 220

Asp Gly Gly Lys Phe Asp Trp Arg Ala Ser Gly Lys Phe Pro Asp Phe
 225 230 235 240

Thr Thr Pro Asp Lys Ser Tyr Asn Gly Leu Ile Tyr Ala Asp Leu Gly
 245 250 255

Ala Pro Ala Phe Ala Leu Lys Ala Arg Val Gln Leu Leu Arg Asn Thr
 260 265 270

Gly Ala Thr Leu Ser Pro Gln Ser Ala Phe Tyr Phe Leu Gln Gly Leu
 275 280 285

Glu Ser Leu Ser Leu Arg Val Gln Lys His Val Asp Asn Thr Arg Lys
 290 295 300

Val Val Glu Phe Leu Lys Asn His Pro Lys Val Ser Trp Ile Asn Tyr
 305 310 315 320

Pro Glu Leu Glu Glu Ser Pro Tyr Lys Glu Leu Ala Asn Lys Tyr Leu
 325 330 335

Pro Lys Gly Ala Gly Ser Ile Phe Thr Phe Gly Ile Lys Gly Gly Leu
 340 345 350

Glu Ala Gly Lys Arg Phe Ile Asn Ser Val Lys Leu Phe Ser Leu Leu
 355 360 365

Ala Asn Val Ala Asp Ala Lys Ser Leu Val Ile His Pro Ser Ser Thr
 370 375 380

Thr His Ala Glu Leu Asn Glu Glu Glu Gln Lys Ala Ala Gly Val Thr
 385 390 395 400

Pro Asp Met Ile Arg Leu Ser Ile Gly Val Glu Asp Ala Glu Asp Leu
 405 410 415

Ile Trp Asp Leu Asn Gln Ala Leu Glu Gln Ala
 420 425

<210> 7

<211> 1293

<212> DNA

<213> Bacillus halodurans

<220>

<221> CDS

<222> (1) .. (1293)

<223>

<400> 7
atg aat cat gaa aac caa tgg cag tta gaa aca aag gcc gtt cat tca 48
Met Asn His Glu Asn Gln Trp Gln Leu Glu Thr Lys Ala Val His Ser
1 5 10 15

gga cag gag atc gat ccg aca acg ttg tcg cga gcc gtc cca ttg tac 96
Gly Gln Glu Ile Asp Pro Thr Thr Leu Ser Arg Ala Val Pro Leu Tyr
20 25 30

caa acg acg tcc tac gga ttt aaa gat aca gac cat gcg gcg aat tta 144
Gln Thr Thr Ser Tyr Gly Phe Lys Asp Thr Asp His Ala Ala Asn Leu
35 40 45

ttt tca cta agt gaa ttt ggc aat atc tat acc cga ttg atg aac cca 192
Phe Ser Leu Ser Glu Phe Gly Asn Ile Tyr Thr Arg Leu Met Asn Pro
50 55 60

acg aca gat gtg ttt gaa aaa cgt gtg gct gcg tta gaa gga gga gcg 240
Thr Thr Asp Val Phe Glu Lys Arg Val Ala Ala Leu Glu Gly Gly Ala
65 70 75 80

gca gct tta gcg acg gcc tca ggg cag gcg gcc att acg tat tcg att 288
Ala Ala Leu Ala Thr Ala Ser Gly Gln Ala Ala Ile Thr Tyr Ser Ile
85 90 95

tta aat att gcg gag gct gga gat gaa atc gtg tcc gct agt agc ctt 336
Leu Asn Ile Ala Glu Ala Gly Asp Glu Ile Val Ser Ala Ser Ser Leu
100 105 110

tac ggc gga acg tat aat tta ttt tcg att acg ttg cca aag cta ggg 384
Tyr Gly Gly Thr Tyr Asn Leu Phe Ser Ile Thr Leu Pro Lys Leu Gly
115 120 125

gta aac gtc cgt ttc gtt gat cca tcg gac cca gaa aac ttc aaa gca 432
Val Asn Val Arg Phe Val Asp Pro Ser Asp Pro Glu Asn Phe Lys Ala
130 135 140

gcg att act gaa aag acg aaa gcc att ttc gct gag tcg att gga aac 480
Ala Ile Thr Glu Lys Thr Lys Ala Ile Phe Ala Glu Ser Ile Gly Asn
145 150 155 160

cct aag gga gac gtg tta gat att gaa gcg gtg gcg aaa gtt gca cac 528
Pro Lys Gly Asp Val Leu Asp Ile Glu Ala Val Ala Lys Val Ala His
165 170 175

gat cat cac ctt ccc ctc att gtc gat aac acg ttt cca agc cca tat 576
Asp His His Leu Pro Leu Ile Val Asp Asn Thr Phe Pro Ser Pro Tyr
180 185 190

ttg ctt caa ccg ata aag cac ggc gca gac att gtt gtg cat tca gca 624
Leu Leu Gln Pro Ile Lys His Gly Ala Asp Ile Val Val His Ser Ala
195 200 205

aca aaa ttt atc ggt ggt cat ggg acg tcg ata gga ggg atc att gtc 672
Thr Lys Phe Ile Gly Gly His Gly Thr Ser Ile Gly Gly Ile Ile Val
210 215 220

gat gga ggg acg ttt gat tgg gcg aaa acg gat cga tat cca ggg cta 720
Asp Gly Gly Thr Phe Asp Trp Ala Lys Thr Asp Arg Tyr Pro Gly Leu
225 230 235 240

aca aca cct gat ccg agt tac cac ggt gtt gta tat aca gat gcg gtc	768
Thr Thr Pro Asp Pro Ser Tyr His Gly Val Val Tyr Thr Asp Ala Val	
245 250 255	
ggt cca att gct tat att att aaa gcg cgt gtt cag cta ttg cgt gac	816
Gly Pro Ile Ala Tyr Ile Ile Lys Ala Arg Val Gln Leu Leu Arg Asp	
260 265 270	
atg ggg gca gcc ata tcg cca ttt aac tcg ttt tta ctg ttg caa ggg	864
Met Gly Ala Ala Ile Ser Pro Phe Asn Ser Phe Leu Leu Leu Gln Gly	
275 280 285	
ttg gaa acg ttg cat tta cgg atg gag aga cat agt gaa aat gcc tac	912
Leu Glu Thr Leu His Leu Arg Met Glu Arg His Ser Glu Asn Ala Tyr	
290 295 300	
aaa gta gca gag ttc ctt gag caa cat caa gcg gtc gaa tcg gtg agc	960
Lys Val Ala Glu Phe Leu Glu Gln His Gln Ala Val Glu Ser Val Ser	
305 310 315 320	
tac tct gga ctg cca tcc cat cca tcc tac cca tta gcg aaa aaa tac	1008
Tyr Ser Gly Leu Pro Ser His Pro Ser Tyr Pro Leu Ala Lys Lys Tyr	
325 330 335	
tta cct aaa ggc caa ggg gct atc tta acg ttc gag gta aag ggc ggc	1056
Leu Pro Lys Gly Gln Gly Ala Ile Leu Thr Phe Glu Val Lys Gly Gly	
340 345 350	
gtt gaa gca gga aag aaa ctc att cat tcg gtc cag cta ttc tcc cac	1104
Val Glu Ala Gly Lys Lys Leu Ile His Ser Val Gln Leu Phe Ser His	
355 360 365	
ctt gcc aac gta ggt gat tca aaa tcg ttg atc atc cat cct gca agc	1152
Leu Ala Asn Val Gly Asp Ser Lys Ser Leu Ile Ile His Pro Ala Ser	
370 375 380	
acg acc cac caa cag ctc tcg gaa gca gaa cag cga gac gca gga gtg	1200
Thr Thr His Gln Gln Leu Ser Glu Ala Glu Gln Arg Asp Ala Gly Val	
385 390 395 400	
aca cct ggg atg atc aga ctt tcg gta gga acc gaa tcg att cat gat	1248
Thr Pro Gly Met Ile Arg Leu Ser Val Gly Thr Glu Ser Ile His Asp	
405 410 415	
att atc acc gat ctc aaa cag gcg att gag gcg agt caa gcg taa	1293
Ile Ile Thr Asp Leu Lys Gln Ala Ile Glu Ala Ser Gln Ala	
420 425 430	

<210> 8

<211> 430

<212> PRT

<213> Bacillus halodurans

<400> 8

Met Asn His Glu Asn Gln Trp Gln Leu Glu Thr Lys Ala Val His Ser
1 5 10 15

Gly Gln Glu Ile Asp Pro Thr Thr Leu Ser Arg Ala Val Pro Leu Tyr
 20 25 30

Gln Thr Thr Ser Tyr Gly Phe Lys Asp Thr Asp His Ala Ala Asn Leu
 35 40 45

Phe Ser Leu Ser Glu Phe Gly Asn Ile Tyr Thr Arg Leu Met Asn Pro
 50 55 60

Thr Thr Asp Val Phe Glu Lys Arg Val Ala Ala Leu Glu Gly Gly Ala
 65 70 75 80

Ala Ala Leu Ala Thr Ala Ser Gly Gln Ala Ala Ile Thr Tyr Ser Ile
 85 90 95

Leu Asn Ile Ala Glu Ala Gly Asp Glu Ile Val Ser Ala Ser Ser Leu
 100 105 110

Tyr Gly Gly Thr Tyr Asn Leu Phe Ser Ile Thr Leu Pro Lys Leu Gly
 115 120 125

Val Asn Val Arg Phe Val Asp Pro Ser Asp Pro Glu Asn Phe Lys Ala
 130 135 140

Ala Ile Thr Glu Lys Thr Lys Ala Ile Phe Ala Glu Ser Ile Gly Asn
 145 150 155 160

Pro Lys Gly Asp Val Leu Asp Ile Glu Ala Val Ala Lys Val Ala His
 165 170 175

Asp His His Leu Pro Leu Ile Val Asp Asn Thr Phe Pro Ser Pro Tyr
 180 185 190

Leu Leu Gln Pro Ile Lys His Gly Ala Asp Ile Val Val His Ser Ala
 195 200 205

Thr Lys Phe Ile Gly Gly His Gly Thr Ser Ile Gly Gly Ile Ile Val
 210 215 220

Asp Gly Gly Thr Phe Asp Trp Ala Lys Thr Asp Arg Tyr Pro Gly Leu
 225 230 235 240

Thr Thr Pro Asp Pro Ser Tyr His Gly Val Val Tyr Thr Asp Ala Val
 245 250 255

Gly Pro Ile Ala Tyr Ile Ile Lys Ala Arg Val Gln Leu Leu Arg Asp

260

265

270

Met Gly Ala Ala Ile Ser Pro Phe Asn Ser Phe Leu Leu Leu Gln Gly
275 280 285

Leu Glu Thr Leu His Leu Arg Met Glu Arg His Ser Glu Asn Ala Tyr
290 295 300

Lys Val Ala Glu Phe Leu Glu Gln His Gln Ala Val Glu Ser Val Ser
305 310 315 320

Tyr Ser Gly Leu Pro Ser His Pro Ser Tyr Pro Leu Ala Lys Lys Tyr
325 330 335

Leu Pro Lys Gly Gln Gly Ala Ile Leu Thr Phe Glu Val Lys Gly Gly
340 345 350

Val Glu Ala Gly Lys Lys Leu Ile His Ser Val Gln Leu Phe Ser His
355 360 365

Leu Ala Asn Val Gly Asp Ser Lys Ser Leu Ile Ile His Pro Ala Ser
370 375 380

Thr Thr His Gln Gln Leu Ser Glu Ala Glu Gln Arg Asp Ala Gly Val
385 390 395 400

Thr Pro Gly Met Ile Arg Leu Ser Val Gly Thr Glu Ser Ile His Asp
405 410 415

Ile Ile Thr Asp Leu Lys Gln Ala Ile Glu Ala Ser Gln Ala
420 425 430

<210> 9

<211> 1203

<212> DNA

<213> Bacillus stearothermophilus

<220>

<221> CDS

<222> (1)..(1203)

<223>

<400> 9
atg tcg tat gta ttc cgc gac agc gag cac gcg gcc aat ttg ttt ggt 48
Met Ser Tyr Val Phe Arg Asp Ser Glu His Ala Ala Asn Leu Phe Gly
1 5 10 15

ttg aaa gag gaa ggt ttt att tat acg cgc att atg aat cca acg aac 96
Leu Lys Glu Glu Gly Phe Ile Tyr Thr Arg Ile Met Asn Pro Thr Asn
20 25 30

gac gtg ttc gaa aag cgg atc gcg gcg ctt gaa ggc ggc att ggg gcg 144
Asp Val Phe Glu Lys Arg Ile Ala Ala Leu Glu Gly Gly Ile Gly Ala
35 40 45

ctc gcg ctg tca tcg ggg cag gcg gcg gtg ttt tat tcg atc atc aac 192
Leu Ala Leu Ser Ser Gly Gln Ala Ala Val Phe Tyr Ser Ile Ile Asn
50 55 60

atc gcc tcg gcg ggc gat gaa atc gtc tcg tct tcg tcc att tac ggc 240
Ile Ala Ser Ala Gly Asp Glu Ile Val Ser Ser Ser Ser Ile Tyr Gly
65 70 75 80

gga acg tac aac ttg ttc gcc cat acg ctg cgc aag ttc ggc att acg 288
Gly Thr Tyr Asn Leu Phe Ala His Thr Leu Arg Lys Phe Gly Ile Thr
85 90 95

gtg aag ttt gtc gat ccg tcc gac ccc gaa aac ttt gag cgg gcg atc 336
Val Lys Phe Val Asp Pro Ser Asp Pro Glu Asn Phe Glu Arg Ala Ile
100 105 110

acc gac aaa acg aaa gcc ttg ttt gcg gaa acg atc ggc aac ccg aaa 384
Thr Asp Lys Thr Lys Ala Leu Phe Ala Glu Thr Ile Gly Asn Pro Lys
115 120 125

aac gat gtg ttg gac att gaa gcg gtg gcc gac atc gcc cat cgc cat 432
Asn Asp Val Leu Asp Ile Glu Ala Val Ala Asp Ile Ala His Arg His
130 135 140

gcc att ccg ctc att gtc gac aac acg gtg gcc agt cca tac tta ttg 480
Ala Ile Pro Leu Ile Val Asp Asn Thr Val Ala Ser Pro Tyr Leu Leu
145 150 155 160

cgg ccg att gaa ttc ggc gcc gat atc gtc gtc cac tca gcg acg aag 528
Arg Pro Ile Glu Phe Gly Ala Asp Ile Val Val His Ser Ala Thr Lys
165 170 175

ttc atc ggc ggg cac ggc aat tcg atc ggc ggt gtg att gtg gac agc 576
Phe Ile Gly Gly His Gly Asn Ser Ile Gly Gly Val Ile Val Asp Ser
180 185 190

ggc aag ttt gac tgg aaa ggg agc ggc aag ttt ccg gag ttc acc gag 624
Gly Lys Phe Asp Trp Lys Gly Ser Gly Lys Phe Pro Glu Phe Thr Glu
195 200 205

cca gac cca agc tac cac ggt ttg gtg tat gtg gac gcc gtc ggc gaa 672
Pro Asp Pro Ser Tyr His Gly Leu Val Tyr Val Asp Ala Val Gly Glu
210 215 220

gcg gcg tac atc acg aaa gcg cgc atc cag ctc ttg cgc gat ttg gga 720
Ala Ala Tyr Ile Thr Lys Ala Arg Ile Gln Leu Leu Arg Asp Leu Gly
225 230 235 240

gcg gcg ttg tcg ccg ttt aat gcg ttt ttg ctt ttg caa ggg ttg gag 768
Ala Ala Leu Ser Pro Phe Asn Ala Phe Leu Leu Leu Gln Gly Leu Glu

245										250					255					
acg	ctc	cat	ttg	cgg	atg	cag	cgc	cat	agc	gaa	aac	gcc	ctt	gcc	gtc		816			
Thr	Leu	His	Leu	Arg	Met	Gln	Arg	His	Ser	Glu	Asn	Ala	Leu	Ala	Val					
			260					265					270							
gcc	aag	ttt	tta	gaa	gag	gaa	gaa	gcg	gtc	gaa	tcg	gtc	aat	tac	cca		864			
Ala	Lys	Phe	Leu	Glu	Glu	Glu	Glu	Ala	Val	Glu	Ser	Val	Asn	Tyr	Pro					
		275					280					285								
ggg	ctt	ccg	agc	cat	ccg	tcg	cat	gaa	ctg	gcg	aaa	aaa	tat	ttg	cca		912			
Gly	Leu	Pro	Ser	His	Pro	Ser	His	Glu	Leu	Ala	Lys	Lys	Tyr	Leu	Pro					
	290					295					300									
aac	ggg	caa	gga	gcg	atc	gtc	acg	ttt	gaa	atc	aaa	ggc	ggc	gtc	gaa		960			
Asn	Gly	Gln	Gly	Ala	Ile	Val	Thr	Phe	Glu	Ile	Lys	Gly	Gly	Val	Glu					
	305				310					315					320					
gcc	ggc	aaa	aaa	ctg	atc	gac	tcg	gtc	aaa	ctg	ttc	tct	cat	ttg	gcc		1008			
Ala	Gly	Lys	Lys	Leu	Ile	Asp	Ser	Val	Lys	Leu	Phe	Ser	His	Leu	Ala					
				325					330					335						
aac	atc	ggc	gat	tcg	aaa	tcg	ctc	atc	atc	cac	ccg	gcc	agc	aca	acg		1056			
Asn	Ile	Gly	Asp	Ser	Lys	Ser	Leu	Ile	Ile	His	Pro	Ala	Ser	Thr	Thr					
			340					345					350							
cac	gag	cag	ctg	agc	cca	gat	gaa	cag	ctg	tcc	gcc	ggc	gtc	acc	cca		1104			
His	Glu	Gln	Leu	Ser	Pro	Asp	Glu	Gln	Leu	Ser	Ala	Gly	Val	Thr	Pro					
			355				360					365								
ggc	ctt	gtg	cgt	ctg	tcc	gtc	ggc	aca	gaa	gcg	atc	gac	gac	att	ttg		1152			
Gly	Leu	Val	Arg	Leu	Ser	Val	Gly	Thr	Glu	Ala	Ile	Asp	Asp	Ile	Leu					
	370					375					380									
gac	gac	ttg	cgc	caa	gcc	att	cgc	caa	agc	cag	acg	gtg	ggg	gtg	aag		1200			
Asp	Asp	Leu	Arg	Gln	Ala	Ile	Arg	Gln	Ser	Gln	Thr	Val	Gly	Val	Lys					
	385				390				395						400					
tag																	1203			

<210> 10

<211> 400

<212> PRT

<213> Bacillus stearothermophilus

<400> 10

Met	Ser	Tyr	Val	Phe	Arg	Asp	Ser	Glu	His	Ala	Ala	Asn	Leu	Phe	Gly
1				5					10					15	

Leu	Lys	Glu	Glu	Gly	Phe	Ile	Tyr	Thr	Arg	Ile	Met	Asn	Pro	Thr	Asn
		20						25					30		

Asp	Val	Phe	Glu	Lys	Arg	Ile	Ala	Ala	Leu	Glu	Gly	Gly	Ile	Gly	Ala
	35						40					45			

Leu Ala Leu Ser Ser Gly Gln Ala Ala Val Phe Tyr Ser Ile Ile Asn
50 55 60

Ile Ala Ser Ala Gly Asp Glu Ile Val Ser Ser Ser Ser Ile Tyr Gly
65 70 75 80

Gly Thr Tyr Asn Leu Phe Ala His Thr Leu Arg Lys Phe Gly Ile Thr
85 90 95

Val Lys Phe Val Asp Pro Ser Asp Pro Glu Asn Phe Glu Arg Ala Ile
100 105 110

Thr Asp Lys Thr Lys Ala Leu Phe Ala Glu Thr Ile Gly Asn Pro Lys
115 120 125

Asn Asp Val Leu Asp Ile Glu Ala Val Ala Asp Ile Ala His Arg His
130 135 140

Ala Ile Pro Leu Ile Val Asp Asn Thr Val Ala Ser Pro Tyr Leu Leu
145 150 155 160

Arg Pro Ile Glu Phe Gly Ala Asp Ile Val Val His Ser Ala Thr Lys
165 170 175

Phe Ile Gly Gly His Gly Asn Ser Ile Gly Gly Val Ile Val Asp Ser
180 185 190

Gly Lys Phe Asp Trp Lys Gly Ser Gly Lys Phe Pro Glu Phe Thr Glu
195 200 205

Pro Asp Pro Ser Tyr His Gly Leu Val Tyr Val Asp Ala Val Gly Glu
210 215 220

Ala Ala Tyr Ile Thr Lys Ala Arg Ile Gln Leu Leu Arg Asp Leu Gly
225 230 235 240

Ala Ala Leu Ser Pro Phe Asn Ala Phe Leu Leu Leu Gln Gly Leu Glu
245 250 255

Thr Leu His Leu Arg Met Gln Arg His Ser Glu Asn Ala Leu Ala Val
260 265 270

Ala Lys Phe Leu Glu Glu Glu Glu Ala Val Glu Ser Val Asn Tyr Pro
275 280 285

Gly Leu Pro Ser His Pro Ser His Glu Leu Ala Lys Lys Tyr Leu Pro
 290 295 300

Asn Gly Gln Gly Ala Ile Val Thr Phe Glu Ile Lys Gly Gly Val Glu
 305 310 315 320

Ala Gly Lys Lys Leu Ile Asp Ser Val Lys Leu Phe Ser His Leu Ala
 325 330 335

Asn Ile Gly Asp Ser Lys Ser Leu Ile Ile His Pro Ala Ser Thr Thr
 340 345 350

His Glu Gln Leu Ser Pro Asp Glu Gln Leu Ser Ala Gly Val Thr Pro
 355 360 365

Gly Leu Val Arg Leu Ser Val Gly Thr Glu Ala Ile Asp Asp Ile Leu
 370 375 380

Asp Asp Leu Arg Gln Ala Ile Arg Gln Ser Gln Thr Val Gly Val Lys
 385 390 395 400

<210> 11

<211> 1290

<212> DNA

<213> Chlorobium tepidum

<220>

<221> CDS

<222> (1) .. (1290)

<223>

<400> 11	
atg agt gag gat aac acc ttc cgg ttc gag acc ttg cag gtt cac gcc	48
Met Ser Glu Asp Asn Thr Phe Arg Phe Glu Thr Leu Gln Val His Ala	
1 5 10 15	
ggg cag gag cct gat ccg gtg acc gga tcg cgc gcc gtg ccc att tac	96
Gly Gln Glu Pro Asp Pro Val Thr Gly Ser Arg Ala Val Pro Ile Tyr	
20 25 30	
cag acc acc tcc tac gtg ttc gag aac gcc gag cac ggc gct gac ctg	144
Gln Thr Thr Ser Tyr Val Phe Glu Asn Ala Glu His Gly Ala Asp Leu	
35 40 45	
ttc gcg ctt cgc aag gcg ggc aat atc tac acg cgc ctg atg aac ccg	192
phe Ala Leu Arg Lys Ala Gly Asn Ile Tyr Thr Arg Leu Met Asn Pro	

50	55	60	
acc acc gac gtg ctc gaa aag cgc atg gcg gcg ctc gaa ggg ggc aag			240
Thr Thr Asp Val Leu Glu Lys Arg Met Ala Ala Leu Glu Gly Gly Lys			
65	70	75	80
gcg gcc ctc ggc gtg gcg agc ggc cac tcg gcg cag ttc atc gct att			288
Ala Ala Leu Gly Val Ala Ser Gly His Ser Ala Gln Phe Ile Ala Ile			
	85	90	95
gcc acc atc tgc cag gct gga gac aac att gtg tca tcg agc tat ctc			336
Ala Thr Ile Cys Gln Ala Gly Asp Asn Ile Val Ser Ser Ser Tyr Leu			
	100	105	110
tac ggc ggc acc tac aac cag ttc aag gtc gcc ttc aag cgc ctc ggc			384
Tyr Gly Gly Thr Tyr Asn Gln Phe Lys Val Ala Phe Lys Arg Leu Gly			
	115	120	125
atc gag gtg agg ttc gtg gat ggc aac gat cag gag gcg ttc cgc aag			432
Ile Glu Val Arg Phe Val Asp Gly Asn Asp Gln Glu Ala Phe Arg Lys			
	130	135	140
gct atc gac gag aac acg aaa gcg ctc tac atg gag tcc agc ggc aat			480
Ala Ile Asp Glu Asn Thr Lys Ala Leu Tyr Met Glu Ser Ser Gly Asn			
	145	150	160
ccg gcg ttc cat gtg ccc gat ttc gac gct atc gcg aag att gcc cgt			528
Pro Ala Phe His Val Pro Asp Phe Asp Ala Ile Ala Lys Ile Ala Arg			
	165	170	175
gag aac ggc att ccg ctg atc gtc gat aac acc ttt ggc tgc gcg ggc			576
Glu Asn Gly Ile Pro Leu Ile Val Asp Asn Thr Phe Gly Cys Ala Gly			
	180	185	190
tat ctc tgc cgt ccc att gat cac ggc gcg tcg atc gtg gtc gag tcg			624
Tyr Leu Cys Arg Pro Ile Asp His Gly Ala Ser Ile Val Val Glu Ser			
	195	200	205
gcc acc aag tgg atc ggc ggg cac ggc acc tcg atg ggc ggc atc atc			672
Ala Thr Lys Trp Ile Gly Gly His Gly Thr Ser Met Gly Gly Ile Ile			
	210	215	220
gtc gat gcc gga acg ttc gac tgg ggc aac ggc aag ttt ccg ctc ttt			720
Val Asp Ala Gly Thr Phe Asp Trp Gly Asn Gly Lys Phe Pro Leu Phe			
	225	230	240
acc gag cca tcg gaa ggc tat cac ggc ctg aaa ttc tac gaa gcg gtc			768
Thr Glu Pro Ser Glu Gly Tyr His Gly Leu Lys Phe Tyr Glu Ala Val			
	245	250	255
ggc gag ctg gcc ttt atc atc cgg gcg cgg gtc gag gga ctg cgg gat			816
Gly Glu Leu Ala Phe Ile Ile Arg Ala Arg Val Glu Gly Leu Arg Asp			
	260	265	270
ttc ggc ccg gcg atc agc ccg ttc aac tcc ttc atg ctg ttg cag gga			864
Phe Gly Pro Ala Ile Ser Pro Phe Asn Ser Phe Met Leu Leu Gln Gly			
	275	280	285
ctt gaa acg ctc tcg ctt cgc gtg cag cgc cac ctc gac aac acg ctt			912
Leu Glu Thr Leu Ser Leu Arg Val Gln Arg His Leu Asp Asn Thr Leu			
	290	295	300
gaa ctg gcc cgc tgg ctc gaa agg cac gat gcg gtt gcg tgg gtg aac			960

Glu Leu Ala Arg Trp Leu Glu Arg His Asp Ala Val Ala Trp Val Asn
 305 310 315 320
 tat cca ggc ctc gaa agc cat ccg aca cac gcc ctg gca aaa aaa tat 1008
 Tyr Pro Gly Leu Glu Ser His Pro Thr His Ala Leu Ala Lys Lys Tyr
 325 330 335
 ctc acg cat ggc ttc ggc tgc gtg ctg act ttc ggc gtg aag ggt ggt 1056
 Leu Thr His Gly Phe Gly Cys Val Leu Thr Phe Gly Val Lys Gly Gly
 340 345 350
 tat gaa aac gcg gtg aag ttc atc gac agc gtg aag ctg gcg agc cac 1104
 Tyr Glu Asn Ala Val Lys Phe Ile Asp Ser Val Lys Leu Ala Ser His
 355 360 365
 ctg gcc aac gtg ggt gat gca aaa acg ctc gtc att cat ccg gca tcg 1152
 Leu Ala Asn Val Gly Asp Ala Lys Thr Leu Val Ile His Pro Ala Ser
 370 375 380
 acg acg cac cag cag ctc agc gcc gag gaa cag gta tcg gcg ggc gtc 1200
 Thr Thr His Gln Gln Leu Ser Ala Glu Glu Gln Val Ser Ala Gly Val
 385 390 395 400
 acc gcc gat atg gtg cgc gtg tcg gtt ggt atc gag cat atc gat gac 1248
 Thr Ala Asp Met Val Arg Val Ser Val Gly Ile Glu His Ile Asp Asp
 405 410 415
 atc aag gct gat ttc agc cag gct ttc gag aat tta gca tga 1290
 Ile Lys Ala Asp Phe Ser Gln Ala Phe Glu Asn Leu Ala
 420 425

<210> 12

<211> 429

<212> PRT

<213> Chlorobium tepidum

<400> 12

Met Ser Glu Asp Asn Thr Phe Arg Phe Glu Thr Leu Gln Val His Ala
 1 5 10 15
 Gly Gln Glu Pro Asp Pro Val Thr Gly Ser Arg Ala Val Pro Ile Tyr
 20 25 30
 Gln Thr Thr Ser Tyr Val Phe Glu Asn Ala Glu His Gly Ala Asp Leu
 35 40 45
 Phe Ala Leu Arg Lys Ala Gly Asn Ile Tyr Thr Arg Leu Met Asn Pro
 50 55 60
 Thr Thr Asp Val Leu Glu Lys Arg Met Ala Ala Leu Glu Gly Gly Lys
 65 70 75 80

Ala Ala Leu Gly Val Ala Ser Gly His Ser Ala Gln Phe Ile Ala Ile
85 90 95

Ala Thr Ile Cys Gln Ala Gly Asp Asn Ile Val Ser Ser Ser Tyr Leu
100 105 110

Tyr Gly Gly Thr Tyr Asn Gln Phe Lys Val Ala Phe Lys Arg Leu Gly
115 120 125

Ile Glu Val Arg Phe Val Asp Gly Asn Asp Gln Glu Ala Phe Arg Lys
130 135 140

Ala Ile Asp Glu Asn Thr Lys Ala Leu Tyr Met Glu Ser Ser Gly Asn
145 150 155 160

Pro Ala Phe His Val Pro Asp Phe Asp Ala Ile Ala Lys Ile Ala Arg
165 170 175

Glu Asn Gly Ile Pro Leu Ile Val Asp Asn Thr Phe Gly Cys Ala Gly
180 185 190

Tyr Leu Cys Arg Pro Ile Asp His Gly Ala Ser Ile Val Val Glu Ser
195 200 205

Ala Thr Lys Trp Ile Gly Gly His Gly Thr Ser Met Gly Gly Ile Ile
210 215 220

Val Asp Ala Gly Thr Phe Asp Trp Gly Asn Gly Lys Phe Pro Leu Phe
225 230 235 240

Thr Glu Pro Ser Glu Gly Tyr His Gly Leu Lys Phe Tyr Glu Ala Val
245 250 255

Gly Glu Leu Ala Phe Ile Ile Arg Ala Arg Val Glu Gly Leu Arg Asp
260 265 270

Phe Gly Pro Ala Ile Ser Pro Phe Asn Ser Phe Met Leu Leu Gln Gly
275 280 285

Leu Glu Thr Leu Ser Leu Arg Val Gln Arg His Leu Asp Asn Thr Leu
290 295 300

Glu Leu Ala Arg Trp Leu Glu Arg His Asp Ala Val Ala Trp Val Asn
305 310 315 320

Tyr Pro Gly Leu Glu Ser His Pro Thr His Ala Leu Ala Lys Lys Tyr
325 330 335

Leu Thr His Gly Phe Gly Cys Val Leu Thr Phe Gly Val Lys Gly Gly
 340 345 350

Tyr Glu Asn Ala Val Lys Phe Ile Asp Ser Val Lys Leu Ala Ser His
 355 360 365

Leu Ala Asn Val Gly Asp Ala Lys Thr Leu Val Ile His Pro Ala Ser
 370 375 380

Thr Thr His Gln Gln Leu Ser Ala Glu Glu Gln Val Ser Ala Gly Val
 385 390 395 400

Thr Ala Asp Met Val Arg Val Ser Val Gly Ile Glu His Ile Asp Asp
 405 410 415

Ile Lys Ala Asp Phe Ser Gln Ala Phe Glu Asn Leu Ala
 420 425

<210> 13

<211> 1281

<212> DNA

<213> Lactococcus lactis

<220>

<221> CDS

<222> (1) .. (1281)

<223>

<400> 13

atg act aat cac aat tat aaa ttc gac act ttg caa gtc cat gca gga 48
 Met Thr Asn His Asn Tyr Lys Phe Asp Thr Leu Gln Val His Ala Gly
 1 5 10 15

caa gtc cct gat cct gtc acg ggt tca cgc gcc gtt ccg ctc tat caa 96
 Gln Val Pro Asp Pro Val Thr Gly Ser Arg Ala Val Pro Leu Tyr Gln
 20 25 30

aca act tct ttc gtt ttt aac aat tca gac cat gcc gaa gct cgt ttt 144
 Thr Thr Ser Phe Val Phe Asn Ser Asp His Ala Glu Ala Arg Phe
 35 40 45

gct tta caa gat cct gga gct att tat tca cgt tta gga aat cca acc 192
 Ala Leu Gln Asp Pro Gly Ala Ile Tyr Ser Arg Leu Gly Asn Pro Thr
 50 55 60

aac gat gtt ttt gaa gca cgc atc gca gct ctt gaa ggt gga agt gca Asn Asp Val Phe Glu Ala Arg Ile Ala Ala Leu Glu Gly Gly Ser Ala 65 70 75 80	240
gcc ctt ggt gtt ggt tct ggc tca gcc gct att acc tat gcc atc ttg Ala Leu Gly Val Gly Ser Gly Ser Ala Ala Ile Thr Tyr Ala Ile Leu 85 90 95	288
aat atc gct aca gtc ggt gat aat att gtt tcc gca agt acc ctt tat Asn Ile Ala Thr Val Gly Asp Asn Ile Val Ser Ala Ser Thr Leu Tyr 100 105 110	336
ggt gga acc tat cac ctt ttt tct ggg act tta cca aaa tat gga att Gly Gly Thr Tyr His Leu Phe Ser Gly Thr Leu Pro Lys Tyr Gly Ile 115 120 125	384
aca act aaa ttt gtc aat cca gat gac ccg aag aat ttt gaa gag gcg Thr Thr Lys Phe Val Asn Pro Asp Asp Pro Lys Asn Phe Glu Glu Ala 130 135 140	432
att gat gaa aaa acc aaa gct att tat tat gaa act ttg ggc aat ccg Ile Asp Glu Lys Thr Lys Ala Ile Tyr Tyr Glu Thr Leu Gly Asn Pro 145 150 155 160	480
gga aat aat gtg att gat tat gat gcc att ggt caa att gct aaa aaa Gly Asn Asn Val Ile Asp Tyr Asp Ala Ile Gly Gln Ile Ala Lys Lys 165 170 175	528
cat gga att ccc gtt att gtt gat gca acg ttt act acc cct gtg acc His Gly Ile Pro Val Ile Val Asp Ala Thr Phe Thr Thr Pro Val Thr 180 185 190	576
ttt aaa cca ttt gaa cat ggt gct aat gta att gtt cat tca gca acg Phe Lys Pro Phe Glu His Gly Ala Asn Val Ile Val His Ser Ala Thr 195 200 205	624
aaa ttc att ggc ggt cat ggt act tct att ggt gga gtc atc gtt gat Lys Phe Ile Gly Gly His Gly Thr Ser Ile Gly Gly Val Ile Val Asp 210 215 220	672
ggc gga aac ttt gat tgg gca aat ggt aat ttt cct gat ttt aca caa Gly Gly Asn Phe Asp Trp Ala Asn Gly Asn Phe Pro Asp Phe Thr Gln 225 230 235 240	720
gct gat gaa agc tac aat ggg att aaa ttt gcc gaa ttg ggt gaa att Ala Asp Glu Ser Tyr Asn Gly Ile Lys Phe Ala Glu Leu Gly Glu Ile 245 250 255	768
gct ttt gtg act cgg gtt aga gct att tta tta cgt gat acg ggt gcg Ala Phe Val Thr Arg Val Arg Ala Ile Leu Leu Arg Asp Thr Gly Ala 260 265 270	816
gct tta tca cct ttt cat tct tgg ctt ttc tta cag ggg cta gaa aca Ala Leu Ser Pro Phe His Ser Trp Leu Phe Leu Gln Gly Leu Glu Thr 275 280 285	864
ctc tca ctc cgg gta gaa cgt cac atc tcc aat act aaa aag att gta Leu Ser Leu Arg Val Glu Arg His Ile Ser Asn Thr Lys Lys Ile Val 290 295 300	912

gaa ttt tta gac aat cat cct aag gtg gaa ctt gtt aac cat cct ctg 960
 Glu Phe Leu Asp Asn His Pro Lys Val Glu Leu Val Asn His Pro Leu
 305 310 315 320

ctt gaa agt aat tcc tat cat gcg ctc tat cag aaa tat tat cca aaa 1008
 Leu Glu Ser Asn Ser Tyr His Ala Leu Tyr Gln Lys Tyr Tyr Pro Lys
 325 330 335

gat gct gga tct atc ttt acc ttt gaa ctc aaa gac aaa gat gag aaa 1056
 Asp Ala Gly Ser Ile Phe Thr Phe Glu Leu Lys Asp Lys Asp Glu Lys
 340 345 350

aaa gcg cgt gat ttg att gat cat ctt gaa att ttc tca ctt cta gcc 1104
 Lys Ala Arg Asp Leu Ile Asp His Leu Glu Ile Phe Ser Leu Leu Ala
 355 360 365

aac gtt gga gat acc aaa tca ttg gcc att cat cct gct tcg acc act 1152
 Asn Val Gly Asp Thr Lys Ser Leu Ala Ile His Pro Ala Ser Thr Thr
 370 375 380

cac cag cag ctg aat gcc gaa gaa ctt gct agt gca ggg att tcc aaa 1200
 His Gln Gln Leu Asn Ala Glu Glu Leu Ala Ser Ala Gly Ile Ser Lys
 385 390 395 400

gga acc att cga tta tcg gtt ggt att gaa gat gta act gac ttg att 1248
 Gly Thr Ile Arg Leu Ser Val Gly Ile Glu Asp Val Thr Asp Leu Ile
 405 410 415

gct gat tta gag caa gca tta gaa aaa ata taa 1281
 Ala Asp Leu Glu Gln Ala Leu Glu Lys Ile
 420 425

<210> 14

<211> 426

<212> PRT

<213> Lactococcus lactis

<400> 14

Met Thr Asn His Asn Tyr Lys Phe Asp Thr Leu Gln Val His Ala Gly
 1 5 10 15

Gln Val Pro Asp Pro Val Thr Gly Ser Arg Ala Val Pro Leu Tyr Gln
 20 25 30

Thr Thr Ser Phe Val Phe Asn Asn Ser Asp His Ala Glu Ala Arg Phe
 35 40 45

Ala Leu Gln Asp Pro Gly Ala Ile Tyr Ser Arg Leu Gly Asn Pro Thr
 50 55 60

Asn Asp Val Phe Glu Ala Arg Ile Ala Ala Leu Glu Gly Gly Ser Ala

65

70

75

80

Ala Leu Gly Val Gly Ser Gly Ser Ala Ala Ile Thr Tyr Ala Ile Leu
85 90 95

Asn Ile Ala Thr Val Gly Asp Asn Ile Val Ser Ala Ser Thr Leu Tyr
100 105 110

Gly Gly Thr Tyr His Leu Phe Ser Gly Thr Leu Pro Lys Tyr Gly Ile
115 120 125

Thr Thr Lys Phe Val Asn Pro Asp Asp Pro Lys Asn Phe Glu Glu Ala
130 135 140

Ile Asp Glu Lys Thr Lys Ala Ile Tyr Tyr Glu Thr Leu Gly Asn Pro
145 150 155 160

Gly Asn Asn Val Ile Asp Tyr Asp Ala Ile Gly Gln Ile Ala Lys Lys
165 170 175

His Gly Ile Pro Val Ile Val Asp Ala Thr Phe Thr Thr Pro Val Thr
180 185 190

Phe Lys Pro Phe Glu His Gly Ala Asn Val Ile Val His Ser Ala Thr
195 200 205

Lys Phe Ile Gly Gly His Gly Thr Ser Ile Gly Gly Val Ile Val Asp
210 215 220

Gly Gly Asn Phe Asp Trp Ala Asn Gly Asn Phe Pro Asp Phe Thr Gln
225 230 235 240

Ala Asp Glu Ser Tyr Asn Gly Ile Lys Phe Ala Glu Leu Gly Glu Ile
245 250 255

Ala Phe Val Thr Arg Val Arg Ala Ile Leu Leu Arg Asp Thr Gly Ala
260 265 270

Ala Leu Ser Pro Phe His Ser Trp Leu Phe Leu Gln Gly Leu Glu Thr
275 280 285

Leu Ser Leu Arg Val Glu Arg His Ile Ser Asn Thr Lys Lys Ile Val
290 295 300

Glu Phe Leu Asp Asn His Pro Lys Val Glu Leu Val Asn His Pro Leu
305 310 315 320

Leu Glu Ser Asn Ser Tyr His Ala Leu Tyr Gln Lys Tyr Tyr Pro Lys
 325 330 335

Asp Ala Gly Ser Ile Phe Thr Phe Glu Leu Lys Asp Lys Asp Glu Lys
 340 345 350

Lys Ala Arg Asp Leu Ile Asp His Leu Glu Ile Phe Ser Leu Leu Ala
 355 360 365

Asn Val Gly Asp Thr Lys Ser Leu Ala Ile His Pro Ala Ser Thr Thr
 370 375 380

His Gln Gln Leu Asn Ala Glu Glu Leu Ala Ser Ala Gly Ile Ser Lys
 385 390 395 400

Gly Thr Ile Arg Leu Ser Val Gly Ile Glu Asp Val Thr Asp Leu Ile
 405 410 415

Ala Asp Leu Glu Gln Ala Leu Glu Lys Ile
 420 425

<210> 15

<211> 1173

<212> DNA

<213> *Synechococcus* sp.

<220>

<221> CDS

<222> (1)..(1173)

<223>

<400> 15
 atg tct cag cgt ttc gaa acc ctc cag ctg cat gcc ggc cag tct cca 48
 Met Ser Gln Arg Phe Glu Thr Leu Gln Leu His Ala Gly Gln Ser Pro
 1 5 10 15

gac tcg gcc acc aat gcc aga gcg gtg ccg att tat cag acc agc tcc 96
 Asp Ser Ala Thr Asn Ala Arg Ala Val Pro Ile Tyr Gln Thr Ser Ser
 20 25 30

tac gtc ttc aac gac gcc gag cac ggc gcc aac ctg ttt gga ctg aag 144
 Tyr Val Phe Asn Asp Ala Glu His Gly Ala Asn Leu Phe Gly Leu Lys
 35 40 45

gaa ttc ggc aac atc tac acc cgt ctg atg aac ccg acg acg gat gtg 192
 Glu Phe Gly Asn Ile Tyr Thr Arg Leu Met Asn Pro Thr Thr Asp Val

50	55	60	
ttc gag aag cgg gtg gcg gcc ctg gaa ggg ggt gtg gcc gcg ctg gcc Phe Glu Lys Arg Val Ala Ala Leu Glu Gly Gly Val Ala Ala Leu Ala 65 70 75 80	240		
aca gcc tcc ggt cag tcg gct cag ttc ctg gcg atc acg aat tgc atg Thr Ala Ser Gly Gln Ser Ala Gln Phe Leu Ala Ile Thr Asn Cys Met 85 90 95	288		
cag gca ggg gat aac ttt gtg tcc acg tcg ttc ctt tac ggc ggc acc Gln Ala Gly Asp Asn Phe Val Ser Thr Ser Phe Leu Tyr Gly Gly Thr 100 105 110	336		
tac aac cag ttc aaa gtg caa ttc ccc cgg ctg ggc atc gac gtg cgc Tyr Asn Gln Phe Lys Val Gln Phe Pro Arg Leu Gly Ile Asp Val Arg 115 120 125	384		
ttc gct gat ggc gac gac gtg gag agc ttt gct gcg cag atc gac gac Phe Ala Asp Gly Asp Asp Val Glu Ser Phe Ala Ala Gln Ile Asp Asp 130 135 140	432		
aaa acc aaa ggc ctc tac gtc gaa gcg atg ggc aat cca cgc ttc aac Lys Thr Lys Gly Leu Tyr Val Glu Ala Met Gly Asn Pro Arg Phe Asn 145 150 155 160	480		
atc ccc gat ttc gag ggc ctc tca gcc ctg gct aaa gag cgc ggc atc Ile Pro Asp Phe Glu Gly Leu Ser Ala Leu Ala Lys Glu Arg Gly Ile 165 170 175	528		
cca ttg atc gtg gac aac acc ttg gga gct tgc ggt gcc ctg atg cgt Pro Leu Ile Val Asp Asn Thr Leu Gly Ala Cys Gly Ala Leu Met Arg 180 185 190	576		
ccg atc gat cat ggc gcg gat gtg gtg gtg gaa agc gcc acc aag tgg Pro Ile Asp His Gly Ala Asp Val Val Val Glu Ser Ala Thr Lys Trp 195 200 205	624		
att ggc ggc cat ggc acc agc ctc ggt ggc gtg atc gtt gat gcc ggc Ile Gly Gly His Gly Thr Ser Leu Gly Gly Val Ile Val Asp Ala Gly 210 215 220	672		
aca ttt aac tgg ggc aat ggc aaa ttc ccg ctg ctg agc caa ccc agt Thr Phe Asn Trp Gly Asn Gly Lys Phe Pro Leu Leu Ser Gln Pro Ser 225 230 235 240	720		
gcg gct tat cac ggc ctt gtg cac tgg gat gcc ttc ggc ttc ggc agc Ala Ala Tyr His Gly Leu Val His Trp Asp Ala Phe Gly Phe Gly Ser 245 250 255	768		
gac gtc tgc aag atg ctg gga gtg ccg gac aac cgc aac gtc gcc ttt Asp Val Cys Lys Met Leu Gly Val Pro Asp Asn Arg Asn Val Ala Phe 260 265 270	816		
gcc ctg cga gcc cgg gtc gag ggt cta cgg gac tgg ggt ccg gcg gtt Ala Leu Arg Ala Arg Val Glu Gly Leu Arg Asp Trp Gly Pro Ala Val 275 280 285	864		
agt ccc ttc aat agc ttc ctg ctg ctg caa ggt cta gaa acc ctc agc Ser Pro Phe Asn Ser Phe Leu Leu Leu Gln Gly Leu Glu Thr Leu Ser 290 295 300	912		
ctg cgg gtg gag cgc cac acg gag aac gcc atg gcg ctg gcc acc tgg	960		

Leu Arg Val Glu Arg His Thr Glu Asn Ala Met Ala Leu Ala Thr Trp
 305 310 315 320
 cta gca acg cac ccc aat gtg gag cat gtg agc tac cca ggc ctg agc 1008
 Leu Ala Thr His Pro Asn Val Glu His Val Ser Tyr Pro Gly Leu Ser
 325 330 335
 agc gat ccg tat cac gca gct gcc aag aaa tac ctg acg ggc cgg ggc 1056
 Ser Asp Pro Tyr His Ala Ala Ala Lys Lys Tyr Leu Thr Gly Arg Gly
 340 345 350
 atg gga tgc atg ctg atg ttc tcg ctc aag ggc ggt tac gac gat gca 1104
 Met Gly Cys Met Leu Met Phe Ser Leu Lys Gly Gly Tyr Asp Asp Ala
 355 360 365
 gtc cgt ttc atc aac agc ctt caa ctg gcc agt cac ctc gcc aat gtg 1152
 Val Arg Phe Ile Asn Ser Leu Gln Leu Ala Ser His Leu Ala Asn Val
 370 375 380
 ggg gat gcc aaa acc tgg tga 1173
 Gly Asp Ala Lys Thr Trp
 385 390

<210> 16

<211> 390

<212> PRT

<213> Synechococcus sp.

<400> 16

Met Ser Gln Arg Phe Glu Thr Leu Gln Leu His Ala Gly Gln Ser Pro
 1 5 10 15

Asp Ser Ala Thr Asn Ala Arg Ala Val Pro Ile Tyr Gln Thr Ser Ser
 20 25 30

Tyr Val Phe Asn Asp Ala Glu His Gly Ala Asn Leu Phe Gly Leu Lys
 35 40 45

Glu Phe Gly Asn Ile Tyr Thr Arg Leu Met Asn Pro Thr Thr Asp Val
 50 55 60

Phe Glu Lys Arg Val Ala Ala Leu Glu Gly Gly Val Ala Ala Leu Ala
 65 70 75 80

Thr Ala Ser Gly Gln Ser Ala Gln Phe Leu Ala Ile Thr Asn Cys Met
 85 90 95

Gln Ala Gly Asp Asn Phe Val Ser Thr Ser Phe Leu Tyr Gly Gly Thr
 100 105 110

Tyr Asn Gln Phe Lys Val Gln Phe Pro Arg Leu Gly Ile Asp Val Arg
115 120 125

Phe Ala Asp Gly Asp Asp Val Glu Ser Phe Ala Ala Gln Ile Asp Asp
130 135 140

Lys Thr Lys Gly Leu Tyr Val Glu Ala Met Gly Asn Pro Arg Phe Asn
145 150 155 160

Ile Pro Asp Phe Glu Gly Leu Ser Ala Leu Ala Lys Glu Arg Gly Ile
165 170 175

Pro Leu Ile Val Asp Asn Thr Leu Gly Ala Cys Gly Ala Leu Met Arg
180 185 190

Pro Ile Asp His Gly Ala Asp Val Val Val Glu Ser Ala Thr Lys Trp
195 200 205

Ile Gly Gly His Gly Thr Ser Leu Gly Gly Val Ile Val Asp Ala Gly
210 215 220

Thr Phe Asn Trp Gly Asn Gly Lys Phe Pro Leu Leu Ser Gln Pro Ser
225 230 235 240

Ala Ala Tyr His Gly Leu Val His Trp Asp Ala Phe Gly Phe Gly Ser
245 250 255

Asp Val Cys Lys Met Leu Gly Val Pro Asp Asn Arg Asn Val Ala Phe
260 265 270

Ala Leu Arg Ala Arg Val Glu Gly Leu Arg Asp Trp Gly Pro Ala Val
275 280 285

Ser Pro Phe Asn Ser Phe Leu Leu Gln Gly Leu Glu Thr Leu Ser
290 295 300

Leu Arg Val Glu Arg His Thr Glu Asn Ala Met Ala Leu Ala Thr Trp
305 310 315 320

Leu Ala Thr His Pro Asn Val Glu His Val Ser Tyr Pro Gly Leu Ser
325 330 335

Ser Asp Pro Tyr His Ala Ala Ala Lys Lys Tyr Leu Thr Gly Arg Gly
340 345 350

Met Gly Cys Met Leu Met Phe Ser Leu Lys Gly Gly Tyr Asp Asp Ala
355 360 365

Val Arg Phe Ile Asn Ser Leu Gln Leu Ala Ser His Leu Ala Asn Val
 370 375 380

Gly Asp Ala Lys Thr Trp
 385 390

<210> 17

<211> 1314

<212> DNA

<213> *Emericella nidulans*

<220>

<221> CDS

<222> (1) .. (1314)

<223>

<400> 17
 atg tcc gac cct tca ccg aaa cgt ttc gag acc ctc cag ctc cat gcg 48
 Met Ser Asp Pro Ser Pro Lys Arg Phe Glu Thr Leu Gln Leu His Ala
 1 5 10 15

ggc cag gag cct gac cct gca act aat tcc cgg gct gtc cca atc tat 96
 Gly Gln Glu Pro Asp Pro Ala Thr Asn Ser Arg Ala Val Pro Ile Tyr
 20 25 30

gcg aca acg tcc tac acc ttc aat gac tcc gca cac ggc gcc agg ctt 144
 Ala Thr Thr Ser Tyr Thr Phe Asn Asp Ser Ala His Gly Ala Arg Leu
 35 40 45

ttt ggc ctc aaa gag ttt ggc aat att tac agc cga att atg aat ccc 192
 Phe Gly Leu Lys Glu Phe Gly Asn Ile Tyr Ser Arg Ile Met Asn Pro
 50 55 60

aca gtc gat gtc ttc gaa aaa cgt att gct gca ctc gag gga ggt gtc 240
 Thr Val Asp Val Phe Glu Lys Arg Ile Ala Ala Leu Glu Gly Gly Val
 65 70 75 80

gct gcg gtg gct gcc tca tct ggc cag gca gcc cag ttc atg gcc atc 288
 Ala Ala Val Ala Ala Ser Ser Gly Gln Ala Ala Gln Phe Met Ala Ile
 85 90 95

tct gct cta gcc cat gct ggt gac aat atc gtt tcc aca agt aat ttg 336
 Ser Ala Leu Ala His Ala Gly Asp Asn Ile Val Ser Thr Ser Asn Leu
 100 105 110

tat ggt ggt aca tac aat cag ttt aag gtc ctt ttc cca cga ctg gga 384
 Tyr Gly Gly Thr Tyr Asn Gln Phe Lys Val Leu Phe Pro Arg Leu Gly
 115 120 125

att acc aca aaa ttc gtg cag gga gac aaa gca gag gac att gcc gcc Ile Thr Thr Lys Phe Val Gln Gly Asp Lys Ala Glu Asp Ile Ala Ala 130 135 140	432
gct atc gat gac cgt acc aag gcc gtc tac gtc gag aca ata gga aac Ala Ile Asp Asp Arg Thr Lys Ala Val Tyr Val Glu Thr Ile Gly Asn 145 150 155 160	480
cct cgc tac aat gtg ccc gac ttt gag gtc att gca aaa gta gcc cat Pro Arg Tyr Asn Val Pro Asp Phe Glu Val Ile Ala Lys Val Ala His 165 170 175	528
gag aag gga att ccc ctt gtg gtt gac aac acc ttc ggt gcc gga ggc Glu Lys Gly Ile Pro Leu Val Val Asp Asn Thr Phe Gly Ala Gly Gly 180 185 190	576
tac ttt gtt cga ccc att gaa cat ggc gcc gac att gtc gtg cac agt Tyr Phe Val Arg Pro Ile Glu His Gly Ala Asp Ile Val Val His Ser 195 200 205	624
gca act aaa tgg att gga ggt cat ggc aca acc atc gga ggc gtt gtc Ala Thr Lys Trp Ile Gly Gly His Gly Thr Thr Ile Gly Gly Val Val 210 215 220	672
gtg gac agc ggc aaa ttc gac tgg ggc aag aac gcc gcg cgg ttt cct Val Asp Ser Gly Lys Phe Asp Trp Gly Lys Asn Ala Ala Arg Phe Pro 225 230 235 240	720
cag ttc acg cag cct tct gaa ggt tac cac ggg ttg aac ttc tgg gag Gln Phe Thr Gln Pro Ser Glu Gly Tyr His Gly Leu Asn Phe Trp Glu 245 250 255	768
acc ttc ggc ccc att gcc ttc gcg att cgt gtc cgg gtc gaa atc ctg Thr Phe Gly Pro Ile Ala Phe Ala Ile Arg Val Arg Val Glu Ile Leu 260 265 270	816
cgc gac ctc ggg tcc gcg ctg aac cct ttc gcc gcg cag cag ctc atc Arg Asp Leu Gly Ser Ala Leu Asn Pro Phe Ala Ala Gln Gln Leu Ile 275 280 285	864
ctg ggt ctg gaa acc cta agc ttg cgc gct gag cgt cat gct tcc aac Leu Gly Leu Glu Thr Leu Ser Leu Arg Ala Glu Arg His Ala Ser Asn 290 295 300	912
gct ctg gcc ctc gcc aac tgg cta aag aag aat gat cac gtc agc tgg Ala Leu Ala Leu Ala Asn Trp Leu Lys Lys Asn Asp His Val Ser Trp 305 310 315 320	960
gtt tct tac gtg ggc cta gaa gag cac tcc agc cac gaa gtt gca aag Val Ser Tyr Val Gly Leu Glu Glu His Ser Ser His Glu Val Ala Lys 325 330 335	1008
aag tac ctc aag cgt ggg ttc ggc ggt gtc cta tcc ttt ggt gtc aag Lys Tyr Leu Lys Arg Gly Phe Gly Gly Val Leu Ser Phe Gly Val Lys 340 345 350	1056
ggt gag gca gcc gtc ggt agc cag gtt gtc gac aac ttt aag ctc atc Gly Glu Ala Ala Val Gly Ser Gln Val Val Asp Asn Phe Lys Leu Ile 355 360 365	1104
tcc aat cta gca aat gtt gga gac tcc aag acc ctc gcg att cac ccc Ser Asn Leu Ala Asn Val Gly Asp Ser Lys Thr Leu Ala Ile His Pro 370 375 380	1152

tgg agc acc act cac gag cag ttg acc gac cag gag cga atc gat tct 1200
 Trp Ser Thr Thr His Glu Gln Leu Thr Asp Gln Glu Arg Ile Asp Ser
 385 390 395 400

ggg gtt acg gaa gat gcc atc cgc atc tct gtc ggc act gag cac atc 1248
 Gly Val Thr Glu Asp Ala Ile Arg Ile Ser Val Gly Thr Glu His Ile
 405 410 415

gac gac atc atc gcc gag ttt gaa cag tca ttt gca gcg acc ttc aaa 1296
 Asp Asp Ile Ile Ala Asp Phe Glu Gln Ser Phe Ala Ala Thr Phe Lys
 420 425 430

gtt gtc cgg agt gct tag 1314
 Val Val Arg Ser Ala
 435

<210> 18

<211> 437

<212> PRT

<213> Emericella nidulans

<400> 18

Met Ser Asp Pro Ser Pro Lys Arg Phe Glu Thr Leu Gln Leu His Ala
 1 5 10 15

Gly Gln Glu Pro Asp Pro Ala Thr Asn Ser Arg Ala Val Pro Ile Tyr
 20 25 30

Ala Thr Thr Ser Tyr Thr Phe Asn Asp Ser Ala His Gly Ala Arg Leu
 35 40 45

Phe Gly Leu Lys Glu Phe Gly Asn Ile Tyr Ser Arg Ile Met Asn Pro
 50 55 60

Thr Val Asp Val Phe Glu Lys Arg Ile Ala Ala Leu Glu Gly Gly Val
 65 70 75 80

Ala Ala Val Ala Ala Ser Ser Gly Gln Ala Ala Gln Phe Met Ala Ile
 85 90 95

Ser Ala Leu Ala His Ala Gly Asp Asn Ile Val Ser Thr Ser Asn Leu
 100 105 110

Tyr Gly Gly Thr Tyr Asn Gln Phe Lys Val Leu Phe Pro Arg Leu Gly
 115 120 125

Ile Thr Thr Lys Phe Val Gln Gly Asp Lys Ala Glu Asp Ile Ala Ala

130

135

140

Ala Ile Asp Asp Arg Thr Lys Ala Val Tyr Val Glu Thr Ile Gly Asn
145 150 155 160

Pro Arg Tyr Asn Val Pro Asp Phe Glu Val Ile Ala Lys Val Ala His
165 170 175

Glu Lys Gly Ile Pro Leu Val Val Asp Asn Thr Phe Gly Ala Gly Gly
180 185 190

Tyr Phe Val Arg Pro Ile Glu His Gly Ala Asp Ile Val Val His Ser
195 200 205

Ala Thr Lys Trp Ile Gly Gly His Gly Thr Thr Ile Gly Gly Val Val
210 215 220

Val Asp Ser Gly Lys Phe Asp Trp Gly Lys Asn Ala Ala Arg Phe Pro
225 230 235 240

Gln Phe Thr Gln Pro Ser Glu Gly Tyr His Gly Leu Asn Phe Trp Glu
245 250 255

Thr Phe Gly Pro Ile Ala Phe Ala Ile Arg Val Arg Val Glu Ile Leu
260 265 270

Arg Asp Leu Gly Ser Ala Leu Asn Pro Phe Ala Ala Gln Gln Leu Ile
275 280 285

Leu Gly Leu Glu Thr Leu Ser Leu Arg Ala Glu Arg His Ala Ser Asn
290 295 300

Ala Leu Ala Leu Ala Asn Trp Leu Lys Lys Asn Asp His Val Ser Trp
305 310 315 320

Val Ser Tyr Val Gly Leu Glu Glu His Ser Ser His Glu Val Ala Lys
325 330 335

Lys Tyr Leu Lys Arg Gly Phe Gly Gly Val Leu Ser Phe Gly Val Lys
340 345 350

Gly Glu Ala Ala Val Gly Ser Gln Val Val Asp Asn Phe Lys Leu Ile
355 360 365

Ser Asn Leu Ala Asn Val Gly Asp Ser Lys Thr Leu Ala Ile His Pro
370 375 380

Trp Ser Thr Thr His Glu Gln Leu Thr Asp Gln Glu Arg Ile Asp Ser
 385 390 395 400

Gly Val Thr Glu Asp Ala Ile Arg Ile Ser Val Gly Thr Glu His Ile
 405 410 415

Asp Asp Ile Ile Ala Asp Phe Glu Gln Ser Phe Ala Ala Thr Phe Lys
 420 425 430

Val Val Arg Ser Ala
 435

<210> 19

<211> 1287

<212> DNA

<213> Bacteroides fragilis

<220>

<221> CDS

<222> (1)..(1287)

<223>

<400> 19
 atg gaa acg aaa aaa tta cat ttt gag act tta caa ctc cat gtt gga 48
 Met Glu Thr Lys Lys Leu His Phe Glu Thr Leu Gln Leu His Val Gly 15
 1 5 10
 cag gag act ccc gac ccg gca acc gat gcg cgt gcc gta cct att tat 96
 Gln Glu Thr Pro Asp Pro Ala Thr Asp Ala Arg Ala Val Pro Ile Tyr 30
 20 25 30
 cag aca act tcc tat gtg ttc cgg gat tcg gcc cat gcc gcc gca cga 144
 Gln Thr Thr Ser Tyr Val Phe Arg Asp Ser Ala His Ala Ala Ala Arg 45
 35 40 45
 ttt gga ttg caa gac cct ggg aat att tat gga cga ctg acc aat tcc 192
 Phe Gly Leu Gln Asp Pro Gly Asn Ile Tyr Gly Arg Leu Thr Asn Ser 60
 50 55 60
 act cag gga gta ttg gag gaa cgc atc gca gca ctt gaa ggg gga gta 240
 Thr Gln Gly Val Leu Glu Glu Arg Ile Ala Ala Leu Glu Gly Gly Val 70 75 80
 65 70 75 80
 ggt ggg ctt gcc gtg gct tcc gga gct gct gcc gtg acc tat gct atc 288
 Gly Gly Leu Ala Val Ala Ser Gly Ala Ala Ala Val Thr Tyr Ala Ile 95
 85 90 95
 gag aat atc acc cgt tcc ggt gat cat att gtg gct gcc aag acc att 336

Glu Asn Ile Thr Arg Ser Gly Asp His Ile Val Ala Ala Lys Thr Ile	
100 105 110	
tat ggg ggc aca tat aac ttg ctg gcg cat act ctg cct gct tat gga	384
Tyr Gly Gly Thr Tyr Asn Leu Leu Ala His Thr Leu Pro Ala Tyr Gly	
115 120 125	
gta acg acc act ttt gta gat ccg tcc gat ctt ttt aat ttc gaa cgg	432
Val Thr Thr Thr Phe Val Asp Pro Ser Asp Leu Phe Asn Phe Glu Arg	
130 135 140	
gcg att cgt gaa aat aca aag gcg ata ttc att gaa act ctg gga aac	480
Ala Ile Arg Glu Asn Thr Lys Ala Ile Phe Ile Glu Thr Leu Gly Asn	
145 150 155 160	
ccc aat tcc aat att atc gat atg gat gcc gta gct gcc att gcc cat	528
Pro Asn Ser Asn Ile Ile Asp Met Asp Ala Val Ala Ala Ile Ala His	
165 170 175	
aaa tat cgg att ccg ctg att gtg gat aat act ttc ggt acg cct tac	576
Lys Tyr Arg Ile Pro Leu Ile Val Asp Asn Thr Phe Gly Thr Pro Tyr	
180 185 190	
ctt atc cgt ccc att gag cac ggg gca gac att gtg gta cat tct gcc	624
Leu Ile Arg Pro Ile Glu His Gly Ala Asp Ile Val Val His Ser Ala	
195 200 205	
aca aaa ttc att ggc gga cac ggc agt tcc ttg gga gga gtt att gtc	672
Thr Lys Phe Ile Gly Gly His Gly Ser Ser Leu Gly Gly Val Ile Val	
210 215 220	
gat tcc ggt aaa ttt gac tgg gtt gct tcc ggt aaa ttc ccg caa ctg	720
Asp Ser Gly Lys Phe Asp Trp Val Ala Ser Gly Lys Phe Pro Gln Leu	
225 230 235 240	
acc gag ccg gat gca agt tat cat ggg gta ccg ttt gtc gat gct gcc	768
Thr Glu Pro Asp Ala Ser Tyr His Gly Val Arg Phe Val Asp Ala Ala	
245 250 255	
ggg gct gct gcc tac att gtc cgt ata cgt gcc gtg ttg ctg cgc gat	816
Gly Ala Ala Ala Tyr Ile Val Arg Ile Arg Ala Val Leu Leu Arg Asp	
260 265 270	
acg ggt gct gcc atc agc ccg ttc aat gct ttt atc ttg ctg caa ggg	864
Thr Gly Ala Ala Ile Ser Pro Phe Asn Ala Phe Ile Leu Leu Gln Gly	
275 280 285	
ttg gag act ttg tct ttg cgt gta gaa ccg cat gtg gcc aat gct ttg	912
Leu Glu Thr Leu Ser Leu Arg Val Glu Arg His Val Ala Asn Ala Leu	
290 295 300	
aag gtt att gat ttt ctg gtg aac cat ccg aag gta gcg gct gtt aat	960
Lys Val Ile Asp Phe Leu Val Asn His Pro Lys Val Ala Ala Val Asn	
305 310 315 320	
cat cca tca ttg ccc ggt cat ccg gat cat gcc atc tat caa cgt tat	1008
His Pro Ser Leu Pro Gly His Pro Asp His Ala Ile Tyr Gln Arg Tyr	
325 330 335	
ttt cct ggc ggg gca ggt tct atc ttc act ttc gag gta aag gga gga	1056
Phe Pro Gly Gly Ala Gly Ser Ile Phe Thr Phe Glu Val Lys Gly Gly	
340 345 350	

acg gag gaa gcg cag aag ttt atc gat agt ctg cag ata ttc tct ttg 1104
 Thr Glu Glu Ala Gln Lys Phe Ile Asp Ser Leu Gln Ile Phe Ser Leu
 355 360 365

ctg gcc aat gtg gcc gat gtg aag tcg ctg gtg att cat ccg ggc act 1152
 Leu Ala Asn Val Ala Asp Val Lys Ser Leu Val Ile His Pro Gly Thr
 370 375 380

acc aca cac tcg cag ttg aat gcg cag gag ctg gag gaa cag ggg att 1200
 Thr Thr His Ser Gln Leu Asn Ala Gln Glu Leu Glu Glu Gln Gly Ile
 385 390 395 400

aaa ccc gga acg gtc aga ctt tcg ata ggt acg gag cat att gag gac 1248
 Lys Pro Gly Thr Val Arg Leu Ser Ile Gly Thr Glu His Ile Glu Asp
 405 410 415

att att gat gac tta cgt cag gca tta gag aaa att taa 1287
 Ile Ile Asp Asp Leu Arg Gln Ala Leu Glu Lys Ile
 420 425

<210> 20

<211> 428

<212> PRT

<213> Bacteroides fragilis

<400> 20

Met Glu Thr Lys Lys Leu His Phe Glu Thr Leu Gln Leu His Val Gly
 1 5 10 15

Gln Glu Thr Pro Asp Pro Ala Thr Asp Ala Arg Ala Val Pro Ile Tyr
 20 25 30

Gln Thr Thr Ser Tyr Val Phe Arg Asp Ser Ala His Ala Ala Ala Arg
 35 40 45

Phe Gly Leu Gln Asp Pro Gly Asn Ile Tyr Gly Arg Leu Thr Asn Ser
 50 55 60

Thr Gln Gly Val Leu Glu Glu Arg Ile Ala Ala Leu Glu Gly Gly Val
 65 70 75 80

Gly Gly Leu Ala Val Ala Ser Gly Ala Ala Ala Val Thr Tyr Ala Ile
 85 90 95

Glu Asn Ile Thr Arg Ser Gly Asp His Ile Val Ala Ala Lys Thr Ile
 100 105 110

Tyr Gly Gly Thr Tyr Asn Leu Leu Ala His Thr Leu Pro Ala Tyr Gly
 115 120 125

Val Thr Thr Thr Phe Val Asp Pro Ser Asp Leu Phe Asn Phe Glu Arg
130 135 140

Ala Ile Arg Glu Asn Thr Lys Ala Ile Phe Ile Glu Thr Leu Gly Asn
145 150 155 160

Pro Asn Ser Asn Ile Ile Asp Met Asp Ala Val Ala Ala Ile Ala His
165 170 175

Lys Tyr Arg Ile Pro Leu Ile Val Asp Asn Thr Phe Gly Thr Pro Tyr
180 185 190

Leu Ile Arg Pro Ile Glu His Gly Ala Asp Ile Val Val His Ser Ala
195 200 205

Thr Lys Phe Ile Gly Gly His Gly Ser Ser Leu Gly Gly Val Ile Val
210 215 220

Asp Ser Gly Lys Phe Asp Trp Val Ala Ser Gly Lys Phe Pro Gln Leu
225 230 235 240

Thr Glu Pro Asp Ala Ser Tyr His Gly Val Arg Phe Val Asp Ala Ala
245 250 255

Gly Ala Ala Ala Tyr Ile Val Arg Ile Arg Ala Val Leu Leu Arg Asp
260 265 270

Thr Gly Ala Ala Ile Ser Pro Phe Asn Ala Phe Ile Leu Leu Gln Gly
275 280 285

Leu Glu Thr Leu Ser Leu Arg Val Glu Arg His Val Ala Asn Ala Leu
290 295 300

Lys Val Ile Asp Phe Leu Val Asn His Pro Lys Val Ala Ala Val Asn
305 310 315 320

His Pro Ser Leu Pro Gly His Pro Asp His Ala Ile Tyr Gln Arg Tyr
325 330 335

Phe Pro Gly Gly Ala Gly Ser Ile Phe Thr Phe Glu Val Lys Gly Gly
340 345 350

Thr Glu Glu Ala Gln Lys Phe Ile Asp Ser Leu Gln Ile Phe Ser Leu
355 360 365

Leu Ala Asn Val Ala Asp Val Lys Ser Leu Val Ile His Pro Gly Thr

370

375

380

Thr Thr His Ser Gln Leu Asn Ala Gln Glu Leu Glu Glu Gln Gly Ile
 385 390 395 400

Lys Pro Gly Thr Val Arg Leu Ser Ile Gly Thr Glu His Ile Glu Asp
 405 410 415

Ile Ile Asp Asp Leu Arg Gln Ala Leu Glu Lys Ile
 420 425

<210> 21

<211> 1278

<212> DNA

<213> Pseudomonas aeruginosa

<220>

<221> CDS

<222> (1) .. (1278)

<223>

<400> 21

atg aaa ctg gaa acc ctg gcc gtc cac gcc ggc tac agc cct gac ccg 48
 Met Lys Leu Glu Thr Leu Ala Val His Ala Gly Tyr Ser Pro Asp Pro
 1 5 10 15

acc acc cgc gcg gtg gcg gtg ccg atc tac cag acc acc tcc tac gcc 96
 Thr Thr Arg Ala Val Ala Val Pro Ile Tyr Gln Thr Thr Ser Tyr Ala
 20 25 30

ttc gac gac acc cag cat ggc gcc gac ctg ttc gac ctg aag gta ccg 144
 Phe Asp Asp Thr Gln His Gly Ala Asp Leu Phe Asp Leu Lys Val Pro
 35 40 45

ggc aac atc tac aca cgg atc atg aac ccc acc aac gac gta ctg gaa 192
 Gly Asn Ile Tyr Thr Arg Ile Met Asn Pro Thr Asn Asp Val Leu Glu
 50 55 60

cag cgc gtc gcg gcg ctg gaa ggc ggg gtc ggg gcg ctg gcg gtg gcc 240
 Gln Arg Val Ala Ala Leu Glu Gly Gly Val Gly Ala Leu Ala Val Ala
 65 70 75 80

tcg ggg atg gcg gcc atc acc tac gcg atc cag acc gtc gcc gag gcc 288
 Ser Gly Met Ala Ala Ile Thr Tyr Ala Ile Gln Thr Val Ala Glu Ala
 85 90 95

ggc gac aac atc gtc tcg gtg gcc aag ctc tac ggc ggc acc tac aac 336
 Gly Asp Asn Ile Val Ser Val Ala Lys Leu Tyr Gly Gly Thr Tyr Asn
 100 105 110

ctg ctg gcc cac acc ctg cca cgc atc ggc atc cag gcg cgc ttc gcc Leu Leu Ala His Thr Leu Pro Arg Ile Gly Ile Gln Ala Arg Phe Ala 115 120 125	384
gcc cac gac gac gtc gcc gcc ctg gaa gcg ctg atc gac gag cgg acc Ala His Asp Asp Val Ala Ala Leu Glu Ala Leu Ile Asp Glu Arg Thr 130 135 140	432
aag gcc gtg ttc tgc gaa acc atc ggc aac ccg gcg ggc aac atc atc Lys Ala Val Phe Cys Glu Thr Ile Gly Asn Pro Ala Gly Asn Ile Ile 145 150 155 160	480
gac ctg cag gca ctg gcc gac gcc gct cac cgc cac ggc gtg cca ctg Asp Leu Gln Ala Leu Ala Asp Ala Ala His Arg His Gly Val Pro Leu 165 170 175	528
atc gtc gac aac acg gta gcc acc ccg gtg ctc tgc cgg ccg ttc gag Ile Val Asp Asn Thr Val Ala Thr Pro Val Leu Cys Arg Pro Phe Glu 180 185 190	576
cac ggc gcc gac atc gtc gtg cac tcg ctg acc aag tac atg ggc ggc His Gly Ala Asp Ile Val Val His Ser Leu Thr Lys Tyr Met Gly Gly 195 200 205	624
cac ggc acc agc atc ggc ggg atc gtg gtc gac tcc ggc aaa ttc gac His Gly Thr Ser Ile Gly Gly Ile Val Val Asp Ser Gly Lys Phe Asp 210 215 220	672
tgg gcg gcg aac aag tcg cgc ttc ccg ctg ctg aac acg ccc gat ccg Trp Ala Ala Asn Lys Ser Arg Phe Pro Leu Leu Asn Thr Pro Asp Pro 225 230 235 240	720
tcc tac cac ggc gtc acc tac acc gag gcc ttc gga ccc gcc gcc ttc Ser Tyr His Gly Val Thr Tyr Thr Glu Ala Phe Gly Pro Ala Ala Phe 245 250 255	768
atc ggc cgc tgc cgg gtg gta ccg ctg cgc aac atg ggc gcg gcg ctc Ile Gly Arg Cys Arg Val Val Pro Leu Arg Asn Met Gly Ala Ala Leu 260 265 270	816
tcg ccg ttc aac gcc ttc ctc atc ctc caa ggc ctg gag acc ctg gcg Ser Pro Phe Asn Ala Phe Leu Ile Leu Gln Gly Leu Glu Thr Leu Ala 275 280 285	864
ctg cgc atg gag cgc cac tgc gac aac gcc ctc gcc gtg gcc cgc tac Leu Arg Met Glu Arg His Cys Asp Asn Ala Leu Ala Val Ala Arg Tyr 290 295 300	912
ctg cag cag cat ccg cag gtg gcc tgg gtg aaa tac gcc ggc ctc gcc Leu Gln Gln His Pro Gln Val Ala Trp Val Lys Tyr Ala Gly Leu Ala 305 310 315 320	960
gac aac ccc gag cac gcc ctg gcc cgg cgc tac ctg ggg ggc cgc ccg Asp Asn Pro Glu His Ala Leu Ala Arg Arg Tyr Leu Gly Gly Arg Pro 325 330 335	1008
gcg gcg atc ctg tct ttc ggc atc cag ggc ggc agc gcc gcc ggc gcg Ala Ala Ile Leu Ser Phe Gly Ile Gln Gly Gly Ser Ala Ala Gly Ala 340 345 350	1056
cgc ttc atc gac gcc ttg aag ctg gtg gtg cgg ctg gtc aac atc ggc Arg Phe Ile Asp Ala Leu Lys Leu Val Val Arg Leu Val Asn Ile Gly	1104

355	360	365	
gac gcc aag tcc ctg gcc tgc cac ccg gcg agc acc acc cac cgc cag			1152
Asp Ala Lys Ser Leu Ala Cys His Pro Ala Ser Thr Thr His Arg Gln			
370	375	380	
ttg aac gcg gag gaa ctg gcc cgc gcc gga gtc tcc gac gac atg gtg			1200
Leu Asn Ala Glu Glu Leu Ala Arg Ala Gly Val Ser Asp Asp Met Val			
385	390	395	400
cgg ctg tcg atc ggc atc gag cac atc gac gac atc ctc gcc gac ctc			1248
Arg Leu Ser Ile Gly Ile Glu His Ile Asp Asp Ile Leu Ala Asp Leu			
405	410	415	
gac cag gcc ctg gcc gcc gcc gca cgc tga			1278
Asp Gln Ala Leu Ala Ala Ala Arg			
420	425		

<210> 22

<211> 425

<212> PRT

<213> Pseudomonas aeruginosa

<400> 22

Met Lys Leu Glu Thr Leu Ala Val His Ala Gly Tyr Ser Pro Asp Pro	
1	15

Thr Thr Arg Ala Val Ala Val Pro Ile Tyr Gln Thr Thr Ser Tyr Ala	
20	30

Phe Asp Asp Thr Gln His Gly Ala Asp Leu Phe Asp Leu Lys Val Pro	
35	45

Gly Asn Ile Tyr Thr Arg Ile Met Asn Pro Thr Asn Asp Val Leu Glu	
50	60

Gln Arg Val Ala Ala Leu Glu Gly Gly Val Gly Ala Leu Ala Val Ala	
65	80

Ser Gly Met Ala Ala Ile Thr Tyr Ala Ile Gln Thr Val Ala Glu Ala	
85	95

Gly Asp Asn Ile Val Ser Val Ala Lys Leu Tyr Gly Gly Thr Tyr Asn	
100	110

Leu Leu Ala His Thr Leu Pro Arg Ile Gly Ile Gln Ala Arg Phe Ala	
115	125

Ala His Asp Asp Val Ala Ala Leu Glu Ala Leu Ile Asp Glu Arg Thr
130 135 140

Lys Ala Val Phe Cys Glu Thr Ile Gly Asn Pro Ala Gly Asn Ile Ile
145 150 155 160

Asp Leu Gln Ala Leu Ala Asp Ala Ala His Arg His Gly Val Pro Leu
165 170 175

Ile Val Asp Asn Thr Val Ala Thr Pro Val Leu Cys Arg Pro Phe Glu
180 185 190

His Gly Ala Asp Ile Val Val His Ser Leu Thr Lys Tyr Met Gly Gly
195 200 205

His Gly Thr Ser Ile Gly Gly Ile Val Val Asp Ser Gly Lys Phe Asp
210 215 220

Trp Ala Ala Asn Lys Ser Arg Phe Pro Leu Leu Asn Thr Pro Asp Pro
225 230 235 240

Ser Tyr His Gly Val Thr Tyr Thr Glu Ala Phe Gly Pro Ala Ala Phe
245 250 255

Ile Gly Arg Cys Arg Val Val Pro Leu Arg Asn Met Gly Ala Ala Leu
260 265 270

Ser Pro Phe Asn Ala Phe Leu Ile Leu Gln Gly Leu Glu Thr Leu Ala
275 280 285

Leu Arg Met Glu Arg His Cys Asp Asn Ala Leu Ala Val Ala Arg Tyr
290 295 300

Leu Gln Gln His Pro Gln Val Ala Trp Val Lys Tyr Ala Gly Leu Ala
305 310 315 320

Asp Asn Pro Glu His Ala Leu Ala Arg Arg Tyr Leu Gly Gly Arg Pro
325 330 335

Ala Ala Ile Leu Ser Phe Gly Ile Gln Gly Gly Ser Ala Ala Gly Ala
340 345 350

Arg Phe Ile Asp Ala Leu Lys Leu Val Val Arg Leu Val Asn Ile Gly
355 360 365

Asp Ala Lys Ser Leu Ala Cys His Pro Ala Ser Thr Thr His Arg Gln
370 375 380

Leu Asn Ala Glu Glu Leu Ala Arg Ala Gly Val Ser Asp Asp Met Val
 385 390 395 400

Arg Leu Ser Ile Gly Ile Glu His Ile Asp Asp Ile Leu Ala Asp Leu
 405 410 415

Asp Gln Ala Leu Ala Ala Ala Arg
 420 425

<210> 23

<211> 1296

<212> DNA

<213> Bordetella bronchiseptica

<220>

<221> CDS

<222> (1) .. (1296)

<223>

<400> 23

atg agc gaa ccg aac caa ccc atc tgg cgg ctg gag acc atc gcc gta	48
Met Ser Glu Pro Asn Gln Pro Ile Trp Arg Leu Glu Thr Ile Ala Val	
1 5 10 15	

cat ggg ggc tac cgg ccc gac ccg acc acg cgc gcg gtg gcg gtg ccg	96
His Gly Gly Tyr Arg Pro Asp Pro Thr Thr Arg Ala Val Ala Val Pro	
20 25 30	

atc tac cag acc gtg gcc tat gcg ttc gac gac acc cag cat ggc gcg	144
Ile Tyr Gln Thr Val Ala Tyr Ala Phe Asp Asp Thr Gln His Gly Ala	
35 40 45	

gac ctg ttc gac ctg aag gtg ccg ggc aat atc tac acc cgc atc atg	192
Asp Leu Phe Asp Leu Lys Val Pro Gly Asn Ile Tyr Thr Arg Ile Met	
50 55 60	

aac ccc acc acc gac gtg ctg gag cag cgc gtg gcg gcg ctg gaa tgc	240
Asn Pro Thr Thr Asp Val Leu Glu Gln Arg Val Ala Ala Leu Glu Cys	
65 70 75 80	

ggc gtg gcc gcg ctg gcg ctg gcc tcc ggc cag gcg gcg gtg acc tat	288
Gly Val Ala Ala Leu Ala Leu Ala Ser Gly Gln Ala Ala Val Thr Tyr	
85 90 95	

gcg atc ctg acc atc gcc gag gcg ggc gac aac atc gtg tcc tcc agc	336
Ala Ile Leu Thr Ile Ala Glu Ala Gly Asp Asn Ile Val Ser Ser Ser	
100 105 110	

acg ctg tat ggc ggc acg tac aac ctg ttc gcc cac acg ctg ccg cag	384
---	-----

Thr	Leu	Tyr	Gly	Gly	Thr	Tyr	Asn	Leu	Phe	Ala	His	Thr	Leu	Pro	Gln		
		115					120					125					
tac	ggc	atc	acg	acc	cgc	ttc	gcc	gat	ccg	cgc	aac	ctg	gct	tcg	ttc	432	
Tyr	Gly	Ile	Thr	Thr	Arg	Phe	Ala	Asp	Pro	Arg	Asn	Leu	Ala	Ser	Phe		
	130					135					140						
gag	gcg	ctg	atc	gac	gag	cgc	acc	aag	gcc	att	ttc	gcc	gag	tcg	gtg	480	
Glu	Ala	Leu	Ile	Asp	Glu	Arg	Thr	Lys	Ala	Ile	Phe	Ala	Glu	Ser	Val		
	145					150				155					160		
ggc	aat	ccg	ctg	ggc	aac	gtc	acc	gac	atc	gcc	gcg	ctg	gcc	gag	atc	528	
Gly	Asn	Pro	Leu	Gly	Asn	Val	Thr	Asp	Ile	Ala	Ala	Leu	Ala	Glu	Ile		
			165						170					175			
gcg	cac	cgc	cat	ggc	gtg	ccg	ctg	atc	gtc	gac	aac	acg	gtg	ccg	tcg	576	
Ala	His	Arg	His	Gly	Val	Pro	Leu	Ile	Val	Asp	Asn	Thr	Val	Pro	Ser		
			180					185					190				
ccc	tac	ctg	ctg	cgc	ccc	atc	gag	cac	ggc	gcc	gac	atc	gtg	gtg	cag	624	
Pro	Tyr	Leu	Leu	Arg	Pro	Ile	Glu	His	Gly	Ala	Asp	Ile	Val	Val	Gln		
	195						200					205					
tcg	ctc	acc	aag	tac	ctg	ggc	ggg	cac	ggc	acc	agc	ctg	ggc	ggg	gcc	672	
Ser	Leu	Thr	Lys	Tyr	Leu	Gly	Gly	His	Gly	Thr	Ser	Leu	Gly	Gly	Ala		
	210					215					220						
atc	atc	gat	tcg	ggc	aag	ttt	ccc	tgg	gcc	gag	cac	aag	gcg	cgc	ttc	720	
Ile	Ile	Asp	Ser	Gly	Lys	Phe	Pro	Trp	Ala	Glu	His	Lys	Ala	Arg	Phe		
	225				230				235						240		
aag	cgc	ctg	aac	gag	ccc	gac	gtg	agc	tac	cac	ggc	gtg	gtc	tac	acc	768	
Lys	Arg	Leu	Asn	Glu	Pro	Asp	Val	Ser	Tyr	His	Gly	Val	Val	Tyr	Thr		
				245					250					255			
gag	gcg	ttc	ggc	gcg	gcg	gcc	tat	atc	ggc	cgc	gcc	cgc	gtg	gtg	ccg	816	
Glu	Ala	Phe	Gly	Ala	Ala	Ala	Tyr	Ile	Gly	Arg	Ala	Arg	Val	Val	Pro		
		260						265					270				
ctg	cgc	aat	acc	ggc	gcg	gcc	att	tcg	ccg	ttc	aac	gcc	ttc	cag	atc	864	
Leu	Arg	Asn	Thr	Gly	Ala	Ala	Ile	Ser	Pro	Phe	Asn	Ala	Phe	Gln	Ile		
		275					280					285					
ctg	cag	ggc	atc	gag	acg	ctg	gcg	ctg	cgc	gtg	gac	cgc	atc	gtc	gag	912	
Leu	Gln	Gly	Ile	Glu	Thr	Leu	Ala	Leu	Arg	Val	Asp	Arg	Ile	Val	Glu		
	290					295					300						
aac	tcg	gtc	aag	gtg	gcc	ggg	ttc	ctg	cgc	gac	cat	ccc	aag	gtc	gaa	960	
Asn	Ser	Val	Lys	Val	Ala	Gly	Phe	Leu	Arg	Asp	His	Pro	Lys	Val	Glu		
	305				310				315						320		
tgg	gtc	aac	tat	gcc	ggc	ctg	ccc	gac	cat	gcc	gac	cat	gcg	ctg	gtg	1008	
Trp	Val	Asn	Tyr	Ala	Gly	Leu	Pro	Asp	His	Ala	Asp	His	Ala	Leu	Val		
				325				330					335				
cgc	aag	tac	atg	ggc	ggc	aag	gcc	ccc	ggc	ctg	ttc	act	ttc	ggc	gtg	1056	
Arg	Lys	Tyr	Met	Gly	Gly	Lys	Ala	Pro	Gly	Leu	Phe	Thr	Phe	Gly	Val		
		340					345					350					
aag	ggc	ggc	cgc	gag	gcc	ggc	gcg	cgc	ttc	cag	gac	gcc	ttg	cag	ctg	1104	
Lys	Gly	Gly	Arg	Glu	Ala	Gly	Ala	Arg	Phe	Gln	Asp	Ala	Leu	Gln	Leu		
		355					360					365					

ttc acc cgc ctg gtg aac atc ggc gac gcc aag tcg ctg gcc acg cac 1152
 Phe Thr Arg Leu Val Asn Ile Gly Asp Ala Lys Ser Leu Ala Thr His
 370 375 380

ccg gct tcc acc acg cac cgc cag ctc aac ccc gaa gag ctc gaa aag 1200
 Pro Ala Ser Thr Thr His Arg Gln Leu Asn Pro Glu Glu Leu Glu Lys
 385 390 395 400

gcc ggc gtg cgc gag gaa acg gtg cgc ctg tcg atc ggg atc gag cat 1248
 Ala Gly Val Arg Glu Glu Thr Val Arg Leu Ser Ile Gly Ile Glu His
 405 410 415

atc gac gac ctg atc gcc gac ctg gaa cag gcg ctg gcg caa gtc tga 1296
 Ile Asp Asp Leu Ile Ala Asp Leu Glu Gln Ala Leu Ala Gln Val
 420 425 430

<210> 24

<211> 431

<212> PRT

<213> Bordetella bronchiseptica

<400> 24

Met Ser Glu Pro Asn Gln Pro Ile Trp Arg Leu Glu Thr Ile Ala Val
 1 5 10 15

His Gly Gly Tyr Arg Pro Asp Pro Thr Thr Arg Ala Val Ala Val Pro
 20 25 30

Ile Tyr Gln Thr Val Ala Tyr Ala Phe Asp Asp Thr Gln His Gly Ala
 35 40 45

Asp Leu Phe Asp Leu Lys Val Pro Gly Asn Ile Tyr Thr Arg Ile Met
 50 55 60

Asn Pro Thr Thr Asp Val Leu Glu Gln Arg Val Ala Ala Leu Glu Cys
 65 70 75 80

Gly Val Ala Ala Leu Ala Leu Ala Ser Gly Gln Ala Ala Val Thr Tyr
 85 90 95

Ala Ile Leu Thr Ile Ala Glu Ala Gly Asp Asn Ile Val Ser Ser Ser
 100 105 110

Thr Leu Tyr Gly Gly Thr Tyr Asn Leu Phe Ala His Thr Leu Pro Gln
 115 120 125

Tyr Gly Ile Thr Thr Arg Phe Ala Asp Pro Arg Asn Leu Ala Ser Phe
 130 135 140

Glu Ala Leu Ile Asp Glu Arg Thr Lys Ala Ile Phe Ala Glu Ser Val
145 150 155 160

Gly Asn Pro Leu Gly Asn Val Thr Asp Ile Ala Ala Leu Ala Glu Ile
165 170 175

Ala His Arg His Gly Val Pro Leu Ile Val Asp Asn Thr Val Pro Ser
180 185 190

Pro Tyr Leu Leu Arg Pro Ile Glu His Gly Ala Asp Ile Val Val Gln
195 200 205

Ser Leu Thr Lys Tyr Leu Gly Gly His Gly Thr Ser Leu Gly Gly Ala
210 215 220

Ile Ile Asp Ser Gly Lys Phe Pro Trp Ala Glu His Lys Ala Arg Phe
225 230 235 240

Lys Arg Leu Asn Glu Pro Asp Val Ser Tyr His Gly Val Val Tyr Thr
245 250 255

Glu Ala Phe Gly Ala Ala Ala Tyr Ile Gly Arg Ala Arg Val Val Pro
260 265 270

Leu Arg Asn Thr Gly Ala Ala Ile Ser Pro Phe Asn Ala Phe Gln Ile
275 280 285

Leu Gln Gly Ile Glu Thr Leu Ala Leu Arg Val Asp Arg Ile Val Glu
290 295 300

Asn Ser Val Lys Val Ala Gly Phe Leu Arg Asp His Pro Lys Val Glu
305 310 315 320

Trp Val Asn Tyr Ala Gly Leu Pro Asp His Ala Asp His Ala Leu Val
325 330 335

Arg Lys Tyr Met Gly Gly Lys Ala Pro Gly Leu Phe Thr Phe Gly Val
340 345 350

Lys Gly Gly Arg Glu Ala Gly Ala Arg Phe Gln Asp Ala Leu Gln Leu
355 360 365

Phe Thr Arg Leu Val Asn Ile Gly Asp Ala Lys Ser Leu Ala Thr His
370 375 380

Pro Ala Ser Thr Thr His Arg Gln Leu Asn Pro Glu Glu Leu Glu Lys

385

390

395

400

Ala Gly Val Arg Glu Glu Thr Val Arg Leu Ser Ile Gly Ile Glu His
405 410 415

Ile Asp Asp Leu Ile Ala Asp Leu Glu Gln Ala Leu Ala Gln Val
420 425 430

<210> 25

<211> 1269

<212> DNA

<213> Nitrosomonas europaea

<220>

<221> CDS

<222> (1)..(1269)

<223>

<400> 25

atg aaa cgg gaa aca ctc gcc att cat ggc ggt ttt gcc ggc gat ccg 48
Met Lys Arg Glu Thr Leu Ala Ile His Gly Gly Phe Ala Gly Asp Pro
1 5 10 15

cag act cat gca gtc gcg gtc ccc att tac cag acc acc agc tac tat 96
Gln Thr His Ala Val Ala Val Pro Ile Tyr Gln Thr Thr Ser Tyr Tyr
20 25 30

ttt gat gat act cag cac ggg gct gat ttg ttt gat ctg aag gtg cag 144
Phe Asp Asp Thr Gln His Gly Ala Asp Leu Phe Asp Leu Lys Val Gln
35 40 45

ggc aac atc tac aca cgc atc atg aac ccg act act gct gtc ctg gaa 192
Gly Asn Ile Tyr Thr Arg Ile Met Asn Pro Thr Thr Ala Val Leu Glu
50 55 60

gaa aga gtg gcg tta ctg gaa gga gga gtg gga gcg ctg gcc atg gct 240
Glu Arg Val Ala Leu Leu Glu Gly Gly Val Gly Ala Leu Ala Met Ala
65 70 75 80

tcc ggc atg gcc gcc att aca gcc tgt gtg cag act ctg gcc agg gcg 288
Ser Gly Met Ala Ala Ile Thr Ala Cys Val Gln Thr Leu Ala Arg Ala
85 90 95

ggc gac aac att atc tcc acc agc cag gtt tac ggt ggc acc tat aat 336
Gly Asp Asn Ile Ile Ser Thr Ser Gln Val Tyr Gly Gly Thr Tyr Asn
100 105 110

ttc ttt tgc cat acg ttg ccc aat ctg ggt att gaa gtt cgc atg gtg 384
Phe Phe Cys His Thr Leu Pro Asn Leu Gly Ile Glu Val Arg Met Val
115 120 125

gat ggt cgt aat ccg gcc gct ttt gcc gat gcc atc gat gac aat acc	432
Asp Gly Arg Asn Pro Ala Ala Phe Ala Asp Ala Ile Asp Asp Asn Thr	
130 135 140	
aga atg att tat tgc gag tcg atc gga aat ccg gcc ggt aat gtg gtg	480
Arg Met Ile Tyr Cys Glu Ser Ile Gly Asn Pro Ala Gly Asn Val Val	
145 150 155 160	
gat atc gcc gca ctg gct gaa gtg gcg cat gca gcg ggc gtg ccg ctg	528
Asp Ile Ala Ala Leu Ala Glu Val Ala His Ala Ala Gly Val Pro Leu	
165 170 175	
gta gtg gac aat acc gta cca acc ccg gtg ctt tgt cgt cct ttc gaa	576
Val Val Asp Asn Thr Val Pro Thr Pro Val Leu Cys Arg Pro Phe Glu	
180 185 190	
cat ggt gcc gat atc gtc gtc cat gcg ctg acc aaa tac atg ggt ggt	624
His Gly Ala Asp Ile Val Val His Ala Leu Thr Lys Tyr Met Gly Gly	
195 200 205	
cac ggc acc agc atc ggc gga atc atc gtg gat tcc ggc aag ttc ccc	672
His Gly Thr Ser Ile Gly Gly Ile Ile Val Asp Ser Gly Lys Phe Pro	
210 215 220	
tgg gaa ggc aac tcg cgt ttt cca caa ttc aac caa cct gat ccc agc	720
Trp Glu Gly Asn Ser Arg Phe Pro Gln Phe Asn Gln Pro Asp Pro Ser	
225 230 235 240	
tat cac ggt gtg gtt tat gtg gat gca ttt ggt ccg gct gcg ttt atc	768
Tyr His Gly Val Val Tyr Val Asp Ala Phe Gly Pro Ala Ala Phe Ile	
245 250 255	
ggc cgt gcg cgt gtg gta ccg ttg cgc aac atg gga gcg gca att tca	816
Gly Arg Ala Arg Val Val Pro Leu Arg Asn Met Gly Ala Ala Ile Ser	
260 265 270	
cct ttc aat tct ttt ctg att ctg caa ggt atc gaa acc ctg ccg ttg	864
Pro Phe Asn Ser Phe Leu Ile Leu Gln Gly Ile Glu Thr Leu Pro Leu	
275 280 285	
agg atg gaa cgg cat tgc acc aat gcg ctg gcg att gca cgt tat ctg	912
Arg Met Glu Arg His Cys Thr Asn Ala Leu Ala Ile Ala Arg Tyr Leu	
290 295 300	
caa agg cat ccc aaa gtc agc tgg gtc aat ttt gcc ggc ctt gaa gat	960
Gln Arg His Pro Lys Val Ser Trp Val Asn Phe Ala Gly Leu Glu Asp	
305 310 315 320	
aac cgt gat tac gca ctg gtg cag aaa tac atg gat ggc ggt att ccc	1008
Asn Arg Asp Tyr Ala Leu Val Gln Lys Tyr Met Asp Gly Gly Ile Pro	
325 330 335	
tca tcg att ctg agt ttt ggc atc aag ggc ggg cgc gag gct tgt gct	1056
Ser Ser Ile Leu Ser Phe Gly Ile Lys Gly Gly Arg Glu Ala Cys Ala	
340 345 350	
cgc ttt atg gac aga ctg atg ctg atc aaa ccg ctg gtc aac atc ggg	1104
Arg Phe Met Asp Arg Leu Met Leu Ile Lys Arg Leu Val Asn Ile Gly	
355 360 365	
gat gcc aaa acg ctg gcc tgc cac ccg gcg acg acc acc cac cgt cag	1152
Asp Ala Lys Thr Leu Ala Cys His Pro Ala Thr Thr Thr His Arg Gln	

370

375

380

ctc aat gat gaa gaa ctg gca aaa gcc ggt gtc agt gct gat ctg gtg 1200
 Leu Asn Asp Glu Glu Leu Ala Lys Ala Gly Val Ser Ala Asp Leu Val
 385 390 395 400

cgt tta tgt gtc ggc atc gag cat att gac gat ctg att gcc gat gta 1248
 Arg Leu Cys Val Gly Ile Glu His Ile Asp Asp Leu Ile Ala Asp Val
 405 410 415

gag cag gct ttc cag gat tag 1269
 Glu Gln Ala Phe Gln Asp
 420

<210> 26

<211> 422

<212> PRT

<213> Nitrosomonas europaea

<400> 26

Met Lys Arg Glu Thr Leu Ala Ile His Gly Gly Phe Ala Gly Asp Pro
 1 5 10 15

Gln Thr His Ala Val Ala Val Pro Ile Tyr Gln Thr Thr Ser Tyr Tyr
 20 25 30

Phe Asp Asp Thr Gln His Gly Ala Asp Leu Phe Asp Leu Lys Val Gln
 35 40 45

Gly Asn Ile Tyr Thr Arg Ile Met Asn Pro Thr Thr Ala Val Leu Glu
 50 55 60

Glu Arg Val Ala Leu Leu Glu Gly Gly Val Gly Ala Leu Ala Met Ala
 65 70 75 80

Ser Gly Met Ala Ala Ile Thr Ala Cys Val Gln Thr Leu Ala Arg Ala
 85 90 95

Gly Asp Asn Ile Ile Ser Thr Ser Gln Val Tyr Gly Gly Thr Tyr Asn
 100 105 110

Phe Phe Cys His Thr Leu Pro Asn Leu Gly Ile Glu Val Arg Met Val
 115 120 125

Asp Gly Arg Asn Pro Ala Ala Phe Ala Asp Ala Ile Asp Asp Asn Thr
 130 135 140

Arg Met Ile Tyr Cys Glu Ser Ile Gly Asn Pro Ala Gly Asn Val Val
145 150 155 160

Asp Ile Ala Ala Leu Ala Glu Val Ala His Ala Ala Gly Val Pro Leu
165 170 175

Val Val Asp Asn Thr Val Pro Thr Pro Val Leu Cys Arg Pro Phe Glu
180 185 190

His Gly Ala Asp Ile Val Val His Ala Leu Thr Lys Tyr Met Gly Gly
195 200 205

His Gly Thr Ser Ile Gly Gly Ile Ile Val Asp Ser Gly Lys Phe Pro
210 215 220

Trp Glu Gly Asn Ser Arg Phe Pro Gln Phe Asn Gln Pro Asp Pro Ser
225 230 235 240

Tyr His Gly Val Val Tyr Val Asp Ala Phe Gly Pro Ala Ala Phe Ile
245 250 255

Gly Arg Ala Arg Val Val Pro Leu Arg Asn Met Gly Ala Ala Ile Ser
260 265 270

Pro Phe Asn Ser Phe Leu Ile Leu Gln Gly Ile Glu Thr Leu Pro Leu
275 280 285

Arg Met Glu Arg His Cys Thr Asn Ala Leu Ala Ile Ala Arg Tyr Leu
290 295 300

Gln Arg His Pro Lys Val Ser Trp Val Asn Phe Ala Gly Leu Glu Asp
305 310 315 320

Asn Arg Asp Tyr Ala Leu Val Gln Lys Tyr Met Asp Gly Gly Ile Pro
325 330 335

Ser Ser Ile Leu Ser Phe Gly Ile Lys Gly Gly Arg Glu Ala Cys Ala
340 345 350

Arg Phe Met Asp Arg Leu Met Leu Ile Lys Arg Leu Val Asn Ile Gly
355 360 365

Asp Ala Lys Thr Leu Ala Cys His Pro Ala Thr Thr Thr His Arg Gln
370 375 380

Leu Asn Asp Glu Glu Leu Ala Lys Ala Gly Val Ser Ala Asp Leu Val
385 390 395 400

Arg Leu Cys Val Gly Ile Glu His Ile Asp Asp Leu Ile Ala Asp Val
 405 410 415

Glu Gln Ala Phe Gln Asp
 420

<210> 27

<211> 1281

<212> DNA

<213> *Sinorhizobium meliloti*

<220>

<221> CDS

<222> (1) .. (1281)

<223>

<400> 27

atg	aaa	gcc	gga	ccc	gga	ttc	agc	acg	ctt	gca	att	cac	gcc	ggg	gcc	48
Met	Lys	Ala	Gly	Pro	Gly	Phe	Ser	Thr	Leu	Ala	Ile	His	Ala	Gly	Ala	
1				5					10					15		

cag	ccc	gat	ccg	acg	acc	ggt	gcg	cgg	gcg	acg	ccg	atc	tat	cag	acg	96
Gln	Pro	Asp	Pro	Thr	Thr	Gly	Ala	Arg	Ala	Thr	Pro	Ile	Tyr	Gln	Thr	
			20					25					30			

acc	agc	ttc	gtc	ttc	aac	gac	acg	gat	cat	gcg	gcc	gca	ctc	ttc	ggc	144
Thr	Ser	Phe	Val	Phe	Asn	Asp	Thr	Asp	His	Ala	Ala	Ala	Leu	Phe	Gly	
		35					40					45				

ctc	cag	caa	ttc	ggc	aat	atc	tat	acc	cgc	atc	atg	aat	ccg	acg	cag	192
Leu	Gln	Gln	Phe	Gly	Asn	Ile	Tyr	Thr	Arg	Ile	Met	Asn	Pro	Thr	Gln	
	50					55					60					

gcg	gtg	ctg	gag	gag	cgg	atc	gcg	gcg	ctc	gaa	ggc	ggg	acc	gcc	ggg	240
Ala	Val	Leu	Glu	Glu	Arg	Ile	Ala	Ala	Leu	Glu	Gly	Gly	Thr	Ala	Gly	
65					70				75						80	

ctg	gcc	gtt	tcc	tcg	ggg	cat	gcg	gcc	cag	ctg	ctg	gtt	ttc	cat	acg	288
Leu	Ala	Val	Ser	Ser	Gly	His	Ala	Ala	Gln	Leu	Leu	Val	Phe	His	Thr	
				85				90						95		

atc	atg	agg	ccg	ggt	gac	aat	ttc	gtt	tcc	gcc	aga	cag	ctt	tac	ggc	336
Ile	Met	Arg	Pro	Gly	Asp	Asn	Phe	Val	Ser	Ala	Arg	Gln	Leu	Tyr	Gly	
			100					105					110			

ggg	tcg	gcc	aat	cag	ttc	ggc	cat	gcc	ttc	aag	gcc	ttc	gac	tgg	cag	384
Gly	Ser	Ala	Asn	Gln	Phe	Gly	His	Ala	Phe	Lys	Ala	Phe	Asp	Trp	Gln	
		115					120					125				

gtc cgc tgg gcc gat tcg gcg gag ccc gaa agc ttc gat gcg cag atc	432
Val Arg Trp Ala Asp Ser Ala Glu Pro Glu Ser Phe Asp Ala Gln Ile	
130 135 140	
gac gaa cgc acc aag gcg atc ttc atc gaa agc ctc gcc aat ccg ggc	480
Asp Glu Arg Thr Lys Ala Ile Phe Ile Glu Ser Leu Ala Asn Pro Gly	
145 150 155 160	
ggc acc ttc gtc gac ata gcc gca atc gct gac gtt gcg cgg cga cac	528
Gly Thr Phe Val Asp Ile Ala Ala Ile Ala Asp Val Ala Arg Arg His	
165 170 175	
gga ctg ccg ctc atc gtc gac aat acg atg gcg acg ccc tat ctg atg	576
Gly Leu Pro Leu Ile Val Asp Asn Thr Met Ala Thr Pro Tyr Leu Met	
180 185 190	
cgg ccg ctc gaa cac ggc gcc gat atc gtc gtc cat tcg ctc acc aag	624
Arg Pro Leu Glu His Gly Ala Asp Ile Val Val His Ser Leu Thr Lys	
195 200 205	
ttc atc ggc ggt cac ggc aat tcg atg ggc ggc atc atc gtc gac ggc	672
Phe Ile Gly Gly His Gly Asn Ser Met Gly Gly Ile Ile Val Asp Gly	
210 215 220	
ggt acg ttc gac tgg tcg aaa tcc ggc aag tat ccg ctg ctg tcg gag	720
Gly Thr Phe Asp Trp Ser Lys Ser Gly Lys Tyr Pro Leu Leu Ser Glu	
225 230 235 240	
ccg agg ccc gaa tat ggc ggc gtc gtc ctg cac cag gcc ttc ggc aac	768
Pro Arg Pro Glu Tyr Gly Gly Val Val Leu His Gln Ala Phe Gly Asn	
245 250 255	
ttc gcc ttc gcc atc gcc gca cgg gta ttg ggt ctg agg gac ttc ggt	816
Phe Ala Phe Ala Ile Ala Ala Arg Val Leu Gly Leu Arg Asp Phe Gly	
260 265 270	
ccg gcc att tcg ccc ttc aac gcc ttc ctg atc cag acc ggc gtc gag	864
Pro Ala Ile Ser Pro Phe Asn Ala Phe Leu Ile Gln Thr Gly Val Glu	
275 280 285	
acg ctg ccg ctg agg atg cag cgc cat tgc gac aac gcg ctg gag gtc	912
Thr Leu Pro Leu Arg Met Gln Arg His Cys Asp Asn Ala Leu Glu Val	
290 295 300	
gcc aaa tgg ctg aag gga cat gaa aag gtc tcc tgg gtc cgc tat tcc	960
Ala Lys Trp Leu Lys Gly His Glu Lys Val Ser Trp Val Arg Tyr Ser	
305 310 315 320	
ggg ctc gaa gac gat ccg aac cac gca ctg cag aaa cgc tac tcg ccg	1008
Gly Leu Glu Asp Asp Pro Asn His Ala Leu Gln Lys Arg Tyr Ser Pro	
325 330 335	
aag ggg gcg gga gcc gtt ttc acc ttc ggg ctc gcg ggc gga tac gag	1056
Lys Gly Ala Gly Ala Val Phe Thr Phe Gly Leu Ala Gly Gly Tyr Glu	
340 345 350	
gcg gga aag cgc ttt gtc gag gca ctg gaa atg ttc tcc cat ctt gcc	1104
Ala Gly Lys Arg Phe Val Glu Ala Leu Glu Met Phe Ser His Leu Ala	
355 360 365	
aat atc ggc gac acg cgt tcg ctc gtc atc cac ccc gca tcg acc acg	1152
Asn Ile Gly Asp Thr Arg Ser Leu Val Ile His Pro Ala Ser Thr Thr	
370 375 380	

cac cgg cag ctc acg ccg gag cag cag gtc gcc gca ggc gcc gga ccc 1200
 His Arg Gln Leu Thr Pro Glu Gln Gln Val Ala Ala Gly Ala Gly Pro
 385 390 395 400

gac gtc atc cgg ttg tcg gtc ggc atc gag gat gtg gcc gac atc att 1248
 Asp Val Ile Arg Leu Ser Val Gly Ile Glu Asp Val Ala Asp Ile Ile
 405 410 415

gcc gat ctc gaa cag gcg ctg ggc aag gcc tga 1281
 Ala Asp Leu Glu Gln Ala Leu Gly Lys Ala
 420 425

<210> 28

<211> 426

<212> PRT

<213> Sinorhizobium meliloti

<400> 28

Met Lys Ala Gly Pro Gly Phe Ser Thr Leu Ala Ile His Ala Gly Ala
 1 5 10 15

Gln Pro Asp Pro Thr Thr Gly Ala Arg Ala Thr Pro Ile Tyr Gln Thr
 20 25 30

Thr Ser Phe Val Phe Asn Asp Thr Asp His Ala Ala Ala Leu Phe Gly
 35 40 45

Leu Gln Gln Phe Gly Asn Ile Tyr Thr Arg Ile Met Asn Pro Thr Gln
 50 55 60

Ala Val Leu Glu Glu Arg Ile Ala Ala Leu Glu Gly Gly Thr Ala Gly
 65 70 75 80

Leu Ala Val Ser Ser Gly His Ala Ala Gln Leu Leu Val Phe His Thr
 85 90 95

Ile Met Arg Pro Gly Asp Asn Phe Val Ser Ala Arg Gln Leu Tyr Gly
 100 105 110

Gly Ser Ala Asn Gln Phe Gly His Ala Phe Lys Ala Phe Asp Trp Gln
 115 120 125

Val Arg Trp Ala Asp Ser Ala Glu Pro Glu Ser Phe Asp Ala Gln Ile
 130 135 140

Asp Glu Arg Thr Lys Ala Ile Phe Ile Glu Ser Leu Ala Asn Pro Gly

145	150	155	160
Gly Thr Phe Val Asp 165	Ile Ala Ala Ile 170	Ala Asp Val Ala Arg Arg 175	His
Gly Leu Pro Leu Ile Val Asp Asn Thr Met Ala Thr Pro Tyr Leu Met 180	185	190	
Arg Pro Leu Glu His Gly Ala Asp Ile Val Val His Ser Leu Thr Lys 195	200	205	
Phe Ile Gly Gly His Gly Asn Ser Met Gly Gly Ile Ile Val Asp Gly 210	215	220	
Gly Thr Phe Asp Trp Ser Lys Ser Gly Lys Tyr Pro Leu Leu Ser Glu 225	230	235	240
Pro Arg Pro Glu Tyr Gly Gly Val Val Leu His Gln Ala Phe Gly Asn 245	250	255	
Phe Ala Phe Ala Ile Ala Ala Arg Val Leu Gly Leu Arg Asp Phe Gly 260	265	270	
Pro Ala Ile Ser Pro Phe Asn Ala Phe Leu Ile Gln Thr Gly Val Glu 275	280	285	
Thr Leu Pro Leu Arg Met Gln Arg His Cys Asp Asn Ala Leu Glu Val 290	295	300	
Ala Lys Trp Leu Lys Gly His Glu Lys Val Ser Trp Val Arg Tyr Ser 305	310	315	320
Gly Leu Glu Asp Asp Pro Asn His Ala Leu Gln Lys Arg Tyr Ser Pro 325	330	335	
Lys Gly Ala Gly Ala Val Phe Thr Phe Gly Leu Ala Gly Gly Tyr Glu 340	345	350	
Ala Gly Lys Arg Phe Val Glu Ala Leu Glu Met Phe Ser His Leu Ala 355	360	365	
Asn Ile Gly Asp Thr Arg Ser Leu Val Ile His Pro Ala Ser Thr Thr 370	375	380	
His Arg Gln Leu Thr Pro Glu Gln Gln Val Ala Ala Gly Ala Gly Pro 385	390	395	400

Asp Val Ile Arg Leu Ser Val Gly Ile Glu Asp Val Ala Asp Ile Ile
 405 410 415

Ala Asp Leu Glu Gln Ala Leu Gly Lys Ala
 420 425

<210> 29

<211> 1293

<212> DNA

<213> Thermotoga maritima

<220>

<221> CDS

<222> (1) .. (1293)

<223>

<400> 29

atg gac tgg aag aaa tac ggt tac aac aca agg gct ctt cac gca ggt	48
Met Asp Trp Lys Lys Tyr Gly Tyr Asn Thr Arg Ala Leu His Ala Gly	
1 5 10 15	

tat gaa cca ccc gag cag gcc aca gga tcg aga gcg gtc cct ata tat	96
Tyr Glu Pro Pro Glu Gln Ala Thr Gly Ser Arg Ala Val Pro Ile Tyr	
20 25 30	

caa acg act tct tac gtt ttc aga gac tct gat cac gcg gcg aga ctc	144
Gln Thr Thr Ser Tyr Val Phe Arg Asp Ser Asp His Ala Ala Arg Leu	
35 40 45	

ttc gca ctg gaa gaa cct ggg ttc atc tat aca agg att gga aat cct	192
Phe Ala Leu Glu Glu Pro Gly Phe Ile Tyr Thr Arg Ile Gly Asn Pro	
50 55 60	

acc gtc tca gtt ctt gaa gaa aga ata gcc gcc ctg gaa gaa ggg gtg	240
Thr Val Ser Val Leu Glu Glu Arg Ile Ala Ala Leu Glu Glu Gly Val	
65 70 75 80	

gga gcc tta gcg gtt gcc agt gga caa gcc gct ata act tac gcc att	288
Gly Ala Leu Ala Val Ala Ser Gly Gln Ala Ala Ile Thr Tyr Ala Ile	
85 90 95	

ttg aac atc gcg ggc cca gga gat gag atc gtc agc ggg agc gcg ctg	336
Leu Asn Ile Ala Gly Pro Gly Asp Glu Ile Val Ser Gly Ser Ala Leu	
100 105 110	

tat ggg gga acg tac aat ctg ttc aga cac act ctc tat aaa aaa tcc	384
Tyr Gly Gly Thr Tyr Asn Leu Phe Arg His Thr Leu Tyr Lys Lys Ser	
115 120 125	

ggc atc atc gtg aag ttt gtg gat gag aca gat cca aag aac ata gaa	432
Gly Ile Ile Val Lys Phe Val Asp Glu Thr Asp Pro Lys Asn Ile Glu	

130

135

140

gag gcc atc acc gag aaa aca aag gcg gtg tac ctt gaa act atc ggg 480
 Glu Ala Ile Thr Glu Lys Thr Lys Ala Val Tyr Leu Glu Thr Ile Gly
 145 150 155 160

aat ccc ggt ctc aca gtg ccg gac ttt gaa gcg ata gcg gag atc gct 528
 Asn Pro Gly Leu Thr Val Pro Asp Phe Glu Ala Ile Ala Glu Ile Ala
 165 170 175

cac aga cac ggt gtt cct ttg ata gtg gac aat acg gta gct ccg tac 576
 His Arg His Gly Val Pro Leu Ile Val Asp Asn Thr Val Ala Pro Tyr
 180 185 190

ata ttc agg ccc ttc gaa cac ggt gcc gac atc gtt gtt tat tcg gcc 624
 Ile Phe Arg Pro Phe Glu His Gly Ala Asp Ile Val Val Tyr Ser Ala
 195 200 205

acg aaa ttc atc gga gga cac gga aca tcg ata ggc ggt ctc atc gta 672
 Thr Lys Phe Ile Gly Gly His Gly Thr Ser Ile Gly Gly Leu Ile Val
 210 215 220

gac agc gga aaa ttc gac tgg acg aac gga aag ttt cca gaa ctc gtg 720
 Asp Ser Gly Lys Phe Asp Trp Thr Asn Gly Lys Phe Pro Glu Leu Val
 225 230 235 240

gaa cca gat ccc agc tac cac ggt gtg agt tat gtg gag acg ttc aaa 768
 Glu Pro Asp Pro Ser Tyr His Gly Val Ser Tyr Val Glu Thr Phe Lys
 245 250 255

gaa gca gcc tac ata gca aaa tgt aga acc cag ctt ttg agg gac ctg 816
 Glu Ala Ala Tyr Ile Ala Lys Cys Arg Thr Gln Leu Leu Arg Asp Leu
 260 265 270

gga agc tgt atg agc ccg ttc aac gcg ttt ctg ttc atc ctc gga ctt 864
 Gly Ser Cys Met Ser Pro Phe Asn Ala Phe Leu Phe Ile Leu Gly Leu
 275 280 285

gaa acc ctc agc ttg agg atg aag aaa cac tgt gaa aac gca ctg aag 912
 Glu Thr Leu Ser Leu Arg Met Lys Lys His Cys Glu Asn Ala Leu Lys
 290 295 300

atc gtt gaa ttt ctg aaa tcg cat ccc gcc gtg agc tgg gtc aac tat 960
 Ile Val Glu Phe Leu Lys Ser His Pro Ala Val Ser Trp Val Asn Tyr
 305 310 315 320

ccg ata gct gaa ggc aat aaa acc aga gaa aat gcg ctg aaa tac ctc 1008
 Pro Ile Ala Glu Gly Asn Lys Thr Arg Glu Asn Ala Leu Lys Tyr Leu
 325 330 335

aaa gaa gga tac ggt gcg att gta acg ttc ggt gtg aaa ggc gga aaa 1056
 Lys Glu Gly Tyr Gly Ala Ile Val Thr Phe Gly Val Lys Gly Gly Lys
 340 345 350

gag gcg gga aag aag ttc ata gac agt ctc aca ctc att tcc cac ctc 1104
 Glu Ala Gly Lys Lys Phe Ile Asp Ser Leu Thr Leu Ile Ser His Leu
 355 360 365

gcc aac att ggt gat gca aga act ctg gct att cat ccc gct tcg aca 1152
 Ala Asn Ile Gly Asp Ala Arg Thr Leu Ala Ile His Pro Ala Ser Thr
 370 375 380

acc cat cag cag ctc acg gaa gaa gag cag ttg aaa acg ggt gtt act 1200

Thr His Gln Gln Leu Thr Glu Glu Glu Gln Leu Lys Thr Gly Val Thr
 385 390 395 400
 ccg gat atg ata aga ttg tct gtt gga ata gaa gat gtg gaa gat atc 1248
 Pro Asp Met Ile Arg Leu Ser Val Gly Ile Glu Asp Val Glu Asp Ile
 405 410 415
 ata gcc gat ctg gat cag gct ctc aga aaa tct cag gag gga tga 1293
 Ile Ala Asp Leu Asp Gln Ala Leu Arg Lys Ser Gln Glu Gly
 420 425 430

<210> 30

<211> 430

<212> PRT

<213> Thermotoga maritima

<400> 30

Met Asp Trp Lys Lys Tyr Gly Tyr Asn Thr Arg Ala Leu His Ala Gly
 1 5 10 15

Tyr Glu Pro Pro Glu Gln Ala Thr Gly Ser Arg Ala Val Pro Ile Tyr
 20 25 30

Gln Thr Thr Ser Tyr Val Phe Arg Asp Ser Asp His Ala Ala Arg Leu
 35 40 45

Phe Ala Leu Glu Glu Pro Gly Phe Ile Tyr Thr Arg Ile Gly Asn Pro
 50 55 60

Thr Val Ser Val Leu Glu Glu Arg Ile Ala Ala Leu Glu Glu Gly Val
 65 70 75 80

Gly Ala Leu Ala Val Ala Ser Gly Gln Ala Ala Ile Thr Tyr Ala Ile
 85 90 95

Leu Asn Ile Ala Gly Pro Gly Asp Glu Ile Val Ser Gly Ser Ala Leu
 100 105 110

Tyr Gly Gly Thr Tyr Asn Leu Phe Arg His Thr Leu Tyr Lys Lys Ser
 115 120 125

Gly Ile Ile Val Lys Phe Val Asp Glu Thr Asp Pro Lys Asn Ile Glu
 130 135 140

Glu Ala Ile Thr Glu Lys Thr Lys Ala Val Tyr Leu Glu Thr Ile Gly
 145 150 155 160

Asn Pro Gly Leu Thr Val Pro Asp Phe Glu Ala Ile Ala Glu Ile Ala
165 170 175

His Arg His Gly Val Pro Leu Ile Val Asp Asn Thr Val Ala Pro Tyr
180 185 190

Ile Phe Arg Pro Phe Glu His Gly Ala Asp Ile Val Val Tyr Ser Ala
195 200 205

Thr Lys Phe Ile Gly Gly His Gly Thr Ser Ile Gly Gly Leu Ile Val
210 215 220

Asp Ser Gly Lys Phe Asp Trp Thr Asn Gly Lys Phe Pro Glu Leu Val
225 230 235 240

Glu Pro Asp Pro Ser Tyr His Gly Val Ser Tyr Val Glu Thr Phe Lys
245 250 255

Glu Ala Ala Tyr Ile Ala Lys Cys Arg Thr Gln Leu Leu Arg Asp Leu
260 265 270

Gly Ser Cys Met Ser Pro Phe Asn Ala Phe Leu Phe Ile Leu Gly Leu
275 280 285

Glu Thr Leu Ser Leu Arg Met Lys Lys His Cys Glu Asn Ala Leu Lys
290 295 300

Ile Val Glu Phe Leu Lys Ser His Pro Ala Val Ser Trp Val Asn Tyr
305 310 315 320

Pro Ile Ala Glu Gly Asn Lys Thr Arg Glu Asn Ala Leu Lys Tyr Leu
325 330 335

Lys Glu Gly Tyr Gly Ala Ile Val Thr Phe Gly Val Lys Gly Gly Lys
340 345 350

Glu Ala Gly Lys Lys Phe Ile Asp Ser Leu Thr Leu Ile Ser His Leu
355 360 365

Ala Asn Ile Gly Asp Ala Arg Thr Leu Ala Ile His Pro Ala Ser Thr
370 375 380

Thr His Gln Gln Leu Thr Glu Glu Glu Gln Leu Lys Thr Gly Val Thr
385 390 395 400

Pro Asp Met Ile Arg Leu Ser Val Gly Ile Glu Asp Val Glu Asp Ile

405

410

415

Ile Ala Asp Leu Asp Gln Ala Leu Arg Lys Ser Gln Glu Gly
 420 425 430

<210> 31

<211> 1314

<212> DNA

<213> Streptococcus mutans

<220>

<221> CDS

<222> (1)..(1314)

<223>

<400> 31

atg	gag	cta	att	aat	aat	aaa	agg	aga	gct	tcc	atg	act	cga	gaa	ttt	48
Met	Glu	Leu	Ile	Asn	Asn	Lys	Arg	Arg	Ala	Ser	Met	Thr	Arg	Glu	Phe	
1				5					10					15		

tct	ttt	gaa	act	tta	caa	tta	cat	gcg	gga	caa	agt	gtt	gat	cct	aca	96
Ser	Phe	Glu	Thr	Leu	Gln	Leu	His	Ala	Gly	Gln	Ser	Val	Asp	Pro	Thr	
			20					25					30			

aca	aaa	tcg	cgt	gca	gta	cca	atc	tat	cag	acg	act	tcc	tat	gtg	ttt	144
Thr	Lys	Ser	Arg	Ala	Val	Pro	Ile	Tyr	Gln	Thr	Thr	Ser	Tyr	Val	Phe	
		35					40					45				

aat	gat	gca	caa	gat	gct	gaa	gat	tct	ttt	gca	ctt	cgt	aca	ccc	ggc	192
Asn	Asp	Ala	Gln	Asp	Ala	Glu	Asp	Ser	Phe	Ala	Leu	Arg	Thr	Pro	Gly	
	50					55					60					

aat	att	tat	acg	cgg	atc	act	aat	ccg	act	aca	gcc	gtt	ttt	gaa	gaa	240
Asn	Ile	Tyr	Thr	Arg	Ile	Thr	Asn	Pro	Thr	Thr	Ala	Val	Phe	Glu	Glu	
	65				70					75				80		

cgg	atg	gcc	gct	ctt	gaa	ggg	ggg	gtc	ggg	gca	ctg	gca	aca	gct	tct	288
Arg	Met	Ala	Ala	Leu	Glu	Gly	Gly	Val	Gly	Ala	Leu	Ala	Thr	Ala	Ser	
			85					90						95		

ggg	atg	gca	gca	gta	act	tat	att	gcc	ttg	gct	ctt	gct	cat	gca	ggg	336
Gly	Met	Ala	Ala	Val	Thr	Tyr	Ile	Ala	Leu	Ala	Leu	Ala	His	Ala	Gly	
			100					105					110			

gat	cat	att	gtg	tca	gca	gcg	aca	gtt	tac	ggg	ggc	act	ttt	aat	ctt	384
Asp	His	Ile	Val	Ser	Ala	Ala	Thr	Val	Tyr	Gly	Gly	Thr	Phe	Asn	Leu	
		115					120					125				

ctt	aag	gaa	act	tta	cct	cgc	tat	ggc	att	act	aca	agt	ttt	gtt	gat	432
Leu	Lys	Glu	Thr	Leu	Pro	Arg	Tyr	Gly	Ile	Thr	Thr	Ser	Phe	Val	Asp	
	130					135					140					

gtt gct aat ttc gct gaa att gaa gcg gct att aca gac aag act aag Val Ala Asn Phe Ala Glu Ile Glu Ala Ala Ile Thr Asp Lys Thr Lys 145 150 155 160	480
ttt att atc gct gaa acg tta gga aat cct ctt gga aat atc gct gat Phe Ile Ile Ala Glu Thr Leu Gly Asn Pro Leu Gly Asn Ile Ala Asp 165 170 175	528
ctt gaa aaa tta gct gag att gcc cat cga cat gct att ccc ttg gtt Leu Glu Lys Leu Ala Glu Ile Ala His Arg His Ala Ile Pro Leu Val 180 185 190	576
att gat aat acc ttt ggt act cct tat ttg ctt aat gtc ttc tct tac Ile Asp Asn Thr Phe Gly Thr Pro Tyr Leu Leu Asn Val Phe Ser Tyr 195 200 205	624
ggt gtt gat att gct gtt cat tct gcc act aaa ttt atc ggt gga cat Gly Val Asp Ile Ala Val His Ser Ala Thr Lys Phe Ile Gly Gly His 210 215 220	672
ggg aca tct att ggc ggt gtc att gtt gat tct gga aac ttt gat tgg Gly Thr Ser Ile Gly Gly Val Ile Val Asp Ser Gly Asn Phe Asp Trp 225 230 235 240	720
gaa aaa tct gga aaa ttc cca caa ttt gta gaa cca gat cct tcc tat Glu Lys Ser Gly Lys Phe Pro Gln Phe Val Glu Pro Asp Pro Ser Tyr 245 250 255	768
cat gac att agt tat aca cgt gat att gga aaa gca gct ttt gta act His Asp Ile Ser Tyr Thr Arg Asp Ile Gly Lys Ala Ala Phe Val Thr 260 265 270	816
gcg gtg cgt acg caa ctg ctg cgt gat aca ggc gcc tgc ctt tca cct Ala Val Arg Thr Gln Leu Leu Arg Asp Thr Gly Ala Cys Leu Ser Pro 275 280 285	864
ttc aat gcc ttt ctt ttg cta caa ggt cta gaa acc tta tca ctt cgt Phe Asn Ala Phe Leu Leu Gln Gly Leu Glu Thr Leu Ser Leu Arg 290 295 300	912
gtt gag cgt cat gtg gaa aat gct aag aaa att gcg tac tat ctg gaa Val Glu Arg His Val Glu Asn Ala Lys Lys Ile Ala Tyr Tyr Leu Glu 305 310 315 320	960
aat cat cct aaa gtc aca aaa gtt aat tat gct agt ttg cca tca agt Asn His Pro Lys Val Thr Lys Val Asn Tyr Ala Ser Leu Pro Ser Ser 325 330 335	1008
cct tat tat gac ttg gct caa aaa tac ttg cca aaa gga gct agt tct Pro Tyr Tyr Asp Leu Ala Gln Lys Tyr Leu Pro Lys Gly Ala Ser Ser 340 345 350	1056
atc ttt act ttt aat gtt gca ggc agt gcg aaa gcc gct cgc gag gtc Ile Phe Thr Phe Asn Val Ala Gly Ser Ala Lys Ala Ala Arg Glu Val 355 360 365	1104
att gac agt ctt gaa atc ttt tct gat ttg gcg aat gtt gct gat gcc Ile Asp Ser Leu Glu Ile Phe Ser Asp Leu Ala Asn Val Ala Asp Ala 370 375 380	1152
aaa tca cta gtt gtt cat ccg gca aca acc act cat ggt caa atg act Lys Ser Leu Val Val His Pro Ala Thr Thr Thr His Gly Gln Met Thr	1200

385	390	395	400	
gaa gaa gat cta cga gct tgc ggt att gaa cct gag caa atc cgt gtt				1248
Glu Glu Asp Leu Arg Ala Cys Gly Ile Glu Pro Glu Gln Ile Arg Val				
405		410	415	
tct att ggt ttg gaa aat gct gat gac tta atc gaa gat ttg cgc cta				1296
Ser Ile Gly Leu Glu Asn Ala Asp Asp Leu Ile Glu Asp Leu Arg Leu				
420	425	430		
gca ctt gaa aaa ata taa				1314
Ala Leu Glu Lys Ile				
435				

<210> 32

<211> 437

<212> PRT

<213> Streptococcus mutans

<400> 32

Met Glu Leu Ile Asn Asn Lys Arg Arg Ala Ser Met Thr Arg Glu Phe
1 5 10 15

Ser Phe Glu Thr Leu Gln Leu His Ala Gly Gln Ser Val Asp Pro Thr
20 25 30

Thr Lys Ser Arg Ala Val Pro Ile Tyr Gln Thr Thr Ser Tyr Val Phe
35 40 45

Asn Asp Ala Gln Asp Ala Glu Asp Ser Phe Ala Leu Arg Thr Pro Gly
50 55 60

Asn Ile Tyr Thr Arg Ile Thr Asn Pro Thr Thr Ala Val Phe Glu Glu
65 70 75 80

Arg Met Ala Ala Leu Glu Gly Gly Val Gly Ala Leu Ala Thr Ala Ser
85 90 95

Gly Met Ala Ala Val Thr Tyr Ile Ala Leu Ala Leu Ala His Ala Gly
100 105 110

Asp His Ile Val Ser Ala Ala Thr Val Tyr Gly Gly Thr Phe Asn Leu
115 120 125

Leu Lys Glu Thr Leu Pro Arg Tyr Gly Ile Thr Thr Ser Phe Val Asp
130 135 140

Val Ala Asn Phe Ala Glu Ile Glu Ala Ala Ile Thr Asp Lys Thr Lys
145 150 155 160

Phe Ile Ile Ala Glu Thr Leu Gly Asn Pro Leu Gly Asn Ile Ala Asp
165 170 175

Leu Glu Lys Leu Ala Glu Ile Ala His Arg His Ala Ile Pro Leu Val
180 185 190

Ile Asp Asn Thr Phe Gly Thr Pro Tyr Leu Leu Asn Val Phe Ser Tyr
195 200 205

Gly Val Asp Ile Ala Val His Ser Ala Thr Lys Phe Ile Gly Gly His
210 215 220

Gly Thr Ser Ile Gly Gly Val Ile Val Asp Ser Gly Asn Phe Asp Trp
225 230 235 240

Glu Lys Ser Gly Lys Phe Pro Gln Phe Val Glu Pro Asp Pro Ser Tyr
245 250 255

His Asp Ile Ser Tyr Thr Arg Asp Ile Gly Lys Ala Ala Phe Val Thr
260 265 270

Ala Val Arg Thr Gln Leu Leu Arg Asp Thr Gly Ala Cys Leu Ser Pro
275 280 285

Phe Asn Ala Phe Leu Leu Leu Gln Gly Leu Glu Thr Leu Ser Leu Arg
290 295 300

Val Glu Arg His Val Glu Asn Ala Lys Lys Ile Ala Tyr Tyr Leu Glu
305 310 315 320

Asn His Pro Lys Val Thr Lys Val Asn Tyr Ala Ser Leu Pro Ser Ser
325 330 335

Pro Tyr Tyr Asp Leu Ala Gln Lys Tyr Leu Pro Lys Gly Ala Ser Ser
340 345 350

Ile Phe Thr Phe Asn Val Ala Gly Ser Ala Lys Ala Ala Arg Glu Val
355 360 365

Ile Asp Ser Leu Glu Ile Phe Ser Asp Leu Ala Asn Val Ala Asp Ala
370 375 380

Lys Ser Leu Val Val His Pro Ala Thr Thr Thr His Gly Gln Met Thr
385 390 395 400

Glu Glu Asp Leu Arg Ala Cys Gly Ile Glu Pro Glu Gln Ile Arg Val
 405 410 415

Ser Ile Gly Leu Glu Asn Ala Asp Asp Leu Ile Glu Asp Leu Arg Leu
 420 425 430

Ala Leu Glu Lys Ile
 435

<210> 33

<211> 1431

<212> DNA

<213> Burkholderia cepacia

<220>

<221> CDS

<222> (1)..(1431)

<223>

<400> 33
 ttg aag cgc cgc acg ccg gtg ata gga tgg ccg cca ctt tca cct ttc 48
 Leu Lys Arg Arg Thr Pro Val Ile Gly Trp Pro Pro Leu Ser Pro Phe
 1 5 10 15
 gcg agg ccg tcc gtg gcc ccg ccg ccc agc atg tcc gcg aac cgt ttc 96
 Ala Arg Pro Ser Val Ala Pro Pro Pro Ser Met Ser Ala Asn Arg Phe
 20 25 30
 gac acg ctt gcg ctg cac gcc ggc gct gct ccc gac ccg acc acc ggc 144
 Asp Thr Leu Ala Leu His Ala Gly Ala Ala Pro Asp Pro Thr Thr Gly
 35 40 45
 gcg cgc gcc acg ccg att tac cag act acc tcg ttt tcg ttc cgc gat 192
 Ala Arg Ala Thr Pro Ile Tyr Gln Thr Thr Ser Phe Ser Phe Arg Asp
 50 55 60
 tcc gac cac gcc gcg gcg ctc ttc aat atg gag cgc gcc ggt cat gtt 240
 Ser Asp His Ala Ala Ala Leu Phe Asn Met Glu Arg Ala Gly His Val
 65 70 75 80
 tat tcg cgc att tcg aac ccg acc gtg gcc gtg ttc gag gaa cgc gtg 288
 Tyr Ser Arg Ile Ser Asn Pro Thr Val Ala Val Phe Glu Glu Arg Val
 85 90 95
 gcc gcg ctg gaa aac ggc gcg ggc gcg atc ggc acg gca agc ggc cag 336
 Ala Ala Leu Glu Asn Gly Ala Gly Ala Ile Gly Thr Ala Ser Gly Gln
 100 105 110
 gcg gcc ctg cat ctg gcc att gcc acg ctg atg ggc gcg ggt tcg cat 384

Ala Ala Leu His Leu Ala Ile Ala Thr Leu Met Gly Ala Gly Ser His	
115 120 125	
atc gtc gcc tcc agc gcg ctg tac ggc ggc tcg cac aat ctg ctg cac	432
Ile Val Ala Ser Ser Ala Leu Tyr Gly Gly Ser His Asn Leu Leu His	
130 135 140	
tac acg ttg cgg cgc ttc ggc atc gag acg act ttc gtc aaa ccc ggc	480
Tyr Thr Leu Arg Arg Phe Gly Ile Glu Thr Thr Phe Val Lys Pro Gly	
145 150 155 160	
gac ctg gac gcg tgg cgc gcc gcg ctg cgc cca aac acg cgg ctg ctg	528
Asp Leu Asp Ala Trp Arg Ala Ala Leu Arg Pro Asn Thr Arg Leu Leu	
165 170 175	
ttc ggc gag acg ctc ggc aat ccg ggg ctc gac gtg ctc gat atc gcc	576
Phe Gly Glu Thr Leu Gly Asn Pro Gly Leu Asp Val Leu Asp Ile Ala	
180 185 190	
gcc gtc gcg cag atc gcg cat gag cac cgc gtg ccg ctg ctg gtc gac	624
Ala Val Ala Gln Ile Ala His Glu His Arg Val Pro Leu Leu Val Asp	
195 200 205	
tcg acc ttc acc aca cct tac ctg ctc aaa ccg ttc gaa cat ggc gcg	672
Ser Thr Phe Thr Thr Pro Tyr Leu Leu Lys Pro Phe Glu His Gly Ala	
210 215 220	
gac ttc gtc tat cac tcg gcc acc aaa ttc ctc ggc ggc cac ggc acg	720
Asp Phe Val Tyr His Ser Ala Thr Lys Phe Leu Gly Gly His Gly Thr	
225 230 235 240	
acg atc ggc ggc gtg ctg gtg gac ggc ggc acg ttc gac ttc gac gcc	768
Thr Ile Gly Gly Val Leu Val Asp Gly Gly Thr Phe Asp Phe Asp Ala	
245 250 255	
tcg ggg cgc ttc ccc gaa ttc acc gaa cct tac gac ggc ttt cac ggc	816
Ser Gly Arg Phe Pro Glu Phe Thr Glu Pro Tyr Asp Gly Phe His Gly	
260 265 270	
atg gtg ttc gcc gag gag agc acc gtc gcg ccg ttt ctg ctg cga gca	864
Met Val Phe Ala Glu Glu Ser Thr Val Ala Pro Phe Leu Leu Arg Ala	
275 280 285	
cgc cgc gag ggg ctg cgc gac ttc ggc gca tgc ctg cat ccg caa gcc	912
Arg Arg Glu Gly Leu Arg Asp Phe Gly Ala Cys Leu His Pro Gln Ala	
290 295 300	
gca tgg caa ctg ctg caa ggc atc gag acg ctg ccg ttg cga atg gaa	960
Ala Trp Gln Leu Leu Gln Gly Ile Glu Thr Leu Pro Leu Arg Met Glu	
305 310 315 320	
cgg cac gtt gcc aac acg cgc cgg gtg gtc gag ttc ctc gcc ggt cac	1008
Arg His Val Ala Asn Thr Arg Arg Val Val Glu Phe Leu Ala Gly His	
325 330 335	
gcc gcg gtc ggg gcc gtc gcc tat ccg gaa ctg ccc acg cac ccc gac	1056
Ala Ala Val Gly Ala Val Ala Tyr Pro Glu Leu Pro Thr His Pro Asp	
340 345 350	
cac gcg ctc gcg aag cgg ctg ctg ccg cgc ggc gcc ggt gcc gtg ttc	1104
His Ala Leu Ala Lys Arg Leu Leu Pro Arg Gly Ala Gly Ala Val Phe	
355 360 365	

agc ttc gat ctg cgc ggc gac cgc gcc gcc gga cgc agc ttt atc gaa 1152
 Ser Phe Asp Leu Arg Gly Asp Arg Ala Ala Gly Arg Ser Phe Ile Glu
 370 375 380

gcg ctc tcg ctg ttc tcg cat ctc gcg aac gtg ggc gac gcg cgc tcg 1200
 Ala Leu Ser Leu Phe Ser His Leu Ala Asn Val Gly Asp Ala Arg Ser
 385 390 395 400

ctc gtg atc cat ccc gcc tcg acc acc cac ttt cgc atg gac gcc gct 1248
 Leu Val Ile His Pro Ala Ser Thr Thr His Phe Arg Met Asp Ala Ala
 405 410 415

gcc ctt gcc gcg gcc ggt atc gcc gaa ggc acg atc cgc ctc tcg atc 1296
 Ala Leu Ala Ala Ala Gly Ile Ala Glu Gly Thr Ile Arg Leu Ser Ile
 420 425 430

ggc ctc gaa gat ccc gac gat ctg atc gac gat ctc aag cgc gcg cta 1344
 Gly Leu Glu Asp Pro Asp Asp Leu Ile Asp Asp Leu Lys Arg Ala Leu
 435 440 445

aag gcc gca cag aaa gcg ggc agt tcg agc gca gcg cac ggc ggc gca 1392
 Lys Ala Ala Gln Lys Ala Gly Ser Ser Ser Ala Ala His Gly Gly Ala
 450 455 460

tcc ggc agt gcc gcc caa ccc cgc ccg gag tcc gca tga 1431
 Ser Gly Ser Ala Ala Gln Pro Arg Pro Glu Ser Ala
 465 470 475

<210> 34

<211> 476

<212> PRT

<213> Burkholderia cepacia

<400> 34

Leu Lys Arg Arg Thr Pro Val Ile Gly Trp Pro Pro Leu Ser Pro Phe
 1 5 10 15

Ala Arg Pro Ser Val Ala Pro Pro Pro Ser Met Ser Ala Asn Arg Phe
 20 25 30

Asp Thr Leu Ala Leu His Ala Gly Ala Ala Pro Asp Pro Thr Thr Gly
 35 40 45

Ala Arg Ala Thr Pro Ile Tyr Gln Thr Thr Ser Phe Ser Phe Arg Asp
 50 55 60

Ser Asp His Ala Ala Ala Leu Phe Asn Met Glu Arg Ala Gly His Val
 65 70 75 80

Tyr Ser Arg Ile Ser Asn Pro Thr Val Ala Val Phe Glu Glu Arg Val
 85 90 95

Ala Ala Leu Glu Asn Gly Ala Gly Ala Ile Gly Thr Ala Ser Gly Gln
100 105 110

Ala Ala Leu His Leu Ala Ile Ala Thr Leu Met Gly Ala Gly Ser His
115 120 125

Ile Val Ala Ser Ser Ala Leu Tyr Gly Gly Ser His Asn Leu Leu His
130 135 140

Tyr Thr Leu Arg Arg Phe Gly Ile Glu Thr Thr Phe Val Lys Pro Gly
145 150 155 160

Asp Leu Asp Ala Trp Arg Ala Ala Leu Arg Pro Asn Thr Arg Leu Leu
165 170 175

Phe Gly Glu Thr Leu Gly Asn Pro Gly Leu Asp Val Leu Asp Ile Ala
180 185 190

Ala Val Ala Gln Ile Ala His Glu His Arg Val Pro Leu Leu Val Asp
195 200 205

Ser Thr Phe Thr Thr Pro Tyr Leu Leu Lys Pro Phe Glu His Gly Ala
210 215 220

Asp Phe Val Tyr His Ser Ala Thr Lys Phe Leu Gly Gly His Gly Thr
225 230 235 240

Thr Ile Gly Gly Val Leu Val Asp Gly Gly Thr Phe Asp Phe Asp Ala
245 250 255

Ser Gly Arg Phe Pro Glu Phe Thr Glu Pro Tyr Asp Gly Phe His Gly
260 265 270

Met Val Phe Ala Glu Glu Ser Thr Val Ala Pro Phe Leu Leu Arg Ala
275 280 285

Arg Arg Glu Gly Leu Arg Asp Phe Gly Ala Cys Leu His Pro Gln Ala
290 295 300

Ala Trp Gln Leu Leu Gln Gly Ile Glu Thr Leu Pro Leu Arg Met Glu
305 310 315 320

Arg His Val Ala Asn Thr Arg Arg Val Val Glu Phe Leu Ala Gly His
325 330 335

Ala Ala Val Gly Ala Val Ala Tyr Pro Glu Leu Pro Thr His Pro Asp

340

345

350

His Ala Leu Ala Lys Arg Leu Leu Pro Arg Gly Ala Gly Ala Val Phe
 355 360 365

Ser Phe Asp Leu Arg Gly Asp Arg Ala Ala Gly Arg Ser Phe Ile Glu
 370 375 380

Ala Leu Ser Leu Phe Ser His Leu Ala Asn Val Gly Asp Ala Arg Ser
 385 390 395 400

Leu Val Ile His Pro Ala Ser Thr Thr His Phe Arg Met Asp Ala Ala
 405 410 415

Ala Leu Ala Ala Ala Gly Ile Ala Glu Gly Thr Ile Arg Leu Ser Ile
 420 425 430

Gly Leu Glu Asp Pro Asp Asp Leu Ile Asp Asp Leu Lys Arg Ala Leu
 435 440 445

Lys Ala Ala Gln Lys Ala Gly Ser Ser Ser Ala Ala His Gly Gly Ala
 450 455 460

Ser Gly Ser Ala Ala Gln Pro Arg Pro Glu Ser Ala
 465 470 475

<210> 35

<211> 1722

<212> DNA

<213> Deinococcus radiodurans

<220>

<221> CDS

<222> (1) .. (1722)

<223>

<400> 35

gtg gcc ttc ccg tgc ggt cag gcg ggg aac aag ata aca agg ccg ggc
 Val Ala Phe Pro Cys Gly Gln Ala Gly Asn Lys Ile Thr Arg Pro Gly
 1 5 10 15

48

caa tgt gtc aac ggg ggc agg gca cgc tca gcc ccg tct aag ttt cgc
 Gln Cys Val Asn Gly Gly Arg Ala Arg Ser Ala Pro Ser Lys Phe Arg
 20 25 30

96

ctt gac ccc tta ccc gcc tcc gcg cta ctt ttt gag gag ctc ccg cag Leu Asp Pro Leu Pro Ala Ser Ala Leu Leu Phe Glu Glu Leu Pro Gln 35 40 45	144
cag gag cca ccc act tca gag cgc ccg aga gac ctg gct cga cga cgg Gln Glu Pro Pro Thr Ser Glu Arg Pro Arg Asp Leu Ala Arg Arg Arg 50 55 60	192
cgc ggc aac cgg acc cca tca cgt cac ggt gcc aag gcc agc ccc ctg Arg Gly Asn Arg Thr Pro Ser Arg His Gly Ala Lys Ala Ser Pro Leu 65 70 75 80	240
ggc gtg tca acg atg agc cgc cgg gcg gga cca agc ggg aag gcc acg Gly Val Ser Thr Met Ser Arg Arg Ala Gly Pro Ser Gly Lys Ala Thr 85 90 95	288
cgg atg acg ata ttc aag tgt ccc ttc tcg att cac agc agg cag ggg Arg Met Thr Ile Phe Lys Cys Pro Phe Ser Ile His Ser Arg Gln Gly 100 105 110	336
gag tgc cgt gac tgg cgc ccc cga acc tgc ttc ccc cga gga gcc gcc Glu Cys Arg Asp Trp Arg Pro Arg Thr Cys Phe Pro Arg Gly Ala Ala 115 120 125	384
acc atg acc gat acc aaa cag ccg cag cct ctg cac ttc gag acc ttg Thr Met Thr Asp Thr Lys Gln Pro Gln Pro Leu His Phe Glu Thr Leu 130 135 140	432
cag gtg cac gcc gga caa cgc ccc gac ccc gtg acc gga gcg cag caa Gln Val His Ala Gly Gln Arg Pro Asp Pro Val Thr Gly Ala Gln Gln 145 150 155 160	480
acg ccc atc tac gcc acc aac tcc tac gtg ttc gag tcg ccc gag cac Thr Pro Ile Tyr Ala Thr Asn Ser Tyr Val Phe Glu Ser Pro Glu His 165 170 175	528
gcc gcc gac ctc ttc ggg ctg cgg caa ttc ggc aac atc tac agc cgc Ala Ala Asp Leu Phe Gly Leu Arg Gln Phe Gly Asn Ile Tyr Ser Arg 180 185 190	576
atc atg aac ccc acc aac gac gtg ttc gag cag cgg gtg gcc gcc ctc Ile Met Asn Pro Thr Asn Asp Val Phe Glu Gln Arg Val Ala Ala Leu 195 200 205	624
gaa ggg ggc gtg ggg gcg ctg tcg gtg tcg agc ggg cac gcg ggg cag Glu Gly Gly Val Gly Ala Leu Ser Val Ser Ser Gly His Ala Gly Gln 210 215 220	672
ctc gtg aca ttg ctc acg ctg gcg cag gcg gga gac aac atc gtc tcg Leu Val Thr Leu Leu Thr Leu Ala Gln Ala Gly Asp Asn Ile Val Ser 225 230 235 240	720
tcg ccc aac ctg tac ggc ggc acc gtc aac cag ttc cgc gtc acg ctc Ser Pro Asn Leu Tyr Gly Gly Thr Val Asn Gln Phe Arg Val Thr Leu 245 250 255	768
aag cgg ctc ggc atc gag gtg cgg ttt acc agc aaa gac gag cgc ccc Lys Arg Leu Gly Ile Glu Val Arg Phe Thr Ser Lys Asp Glu Arg Pro 260 265 270	816
gag gaa ttc gcc gcg ctg atc gac gag cgc acg cgg gcc gta tat ctg Glu Glu Phe Ala Ala Leu Ile Asp Glu Arg Thr Arg Ala Val Tyr Leu	864

275	280	285	
gaa acc atc ggc aac ccg gcg ctg aac att ccc gat ttc gag ggc gtg Glu Thr Ile Gly Asn Pro Ala Leu Asn Ile Pro Asp Phe Glu Gly Val 290 295 300			912
gcg aaa gtc gcg cac gag cac ggc gtc gcg gtg gtc gtg gac aac acc Ala Lys Val Ala His Glu His Gly Val Ala Val Val Val Asp Asn Thr 305 310 315 320			960
ttc ggg gcc ggc gga tac tac tgc cag ccg ctg cgg cac ggc gcc aac Phe Gly Ala Gly Gly Tyr Tyr Cys Gln Pro Leu Arg His Gly Ala Asn 325 330 335			1008
atc gtg ctg cac tcg gcg agc aag tgg atc ggc ggg cac ggc aac ggc Ile Val Leu His Ser Ala Ser Lys Trp Ile Gly Gly His Gly Asn Gly 340 345 350			1056
atc ggc ggg gtc atc gtg gac ggc ggg aac ttc gac tgg ggc agc ggg Ile Gly Gly Val Ile Val Asp Gly Gly Asn Phe Asp Trp Gly Ser Gly 355 360 365			1104
cgg tat ccg ctg atg acc gag ccc tcg ccg agt tat cac ggg ctg aag Arg Tyr Pro Leu Met Thr Glu Pro Ser Pro Ser Tyr His Gly Leu Lys 370 375 380			1152
ttc tgg gag acg ttc ggg gaa ggc aac ggg ctg ggg ctg ccg aac atc Phe Trp Glu Thr Phe Gly Glu Gly Asn Gly Leu Gly Leu Pro Asn Ile 385 390 395 400			1200
gcc ttc atc acc cgc gcc cgc acc gag ggg ctg cgc gac ctg gga acg Ala Phe Ile Thr Arg Ala Arg Thr Glu Gly Leu Arg Asp Leu Gly Thr 405 410 415			1248
acc ctg gcg ccg cag cag gcg tgg cag ttt ctg caa ggc ctt gaa acc Thr Leu Ala Pro Gln Gln Ala Trp Gln Phe Leu Gln Gly Leu Glu Thr 420 425 430			1296
ctg agc ctg cgc gcc gag cgc cac gcc gag aac acc ctg gcg ctg gcg Leu Ser Leu Arg Ala Glu Arg His Ala Glu Asn Thr Leu Ala Leu Ala 435 440 445			1344
cac tgg ctc atc agc cac ccg gac gtg aag cag gtc act tac ccc ggc His Trp Leu Ile Ser His Pro Asp Val Lys Gln Val Thr Tyr Pro Gly 450 455 460			1392
ctg agc aac cac ccc cac tac gac cgg gcg cag acc tac ttg ccg cgc Leu Ser Asn His Pro His Tyr Asp Arg Ala Gln Thr Tyr Leu Pro Arg 465 470 475 480			1440
ggg gcg ggc gcg gtg ctc acc ttc gag ctg cgc ggg ggc cgg gcg gcg Gly Ala Gly Ala Val Leu Thr Phe Glu Leu Arg Gly Gly Arg Ala Ala 485 490 495			1488
ggc gaa gcg ttt att cgc tcg gtc aag ctc gcg cag cac gtc gcc aac Gly Glu Ala Phe Ile Arg Ser Val Lys Leu Ala Gln His Val Ala Asn 500 505 510			1536
gtg ggc gac acc cgc acg ctg gtc att cat ccg gcg agc acc acc cac Val Gly Asp Thr Arg Thr Leu Val Ile His Pro Ala Ser Thr Thr His 515 520 525			1584
agc cag ctc gac gag gtg acg cag acg aac gcc ggg gtc acg ccg ggc			1632

Ser Gln Leu Asp Glu Val Thr Gln Thr Asn Ala Gly Val Thr Pro Gly
 530 535 540

ctc atc cgg gtg tcg gtg ggc atc gag cac gta gac gac atc cgc gag 1680
 Leu Ile Arg Val Ser Val Gly Ile Glu His Val Asp Asp Ile Arg Glu
 545 550 555 560

gac ttc gcg cag gcc ctg gcg agc gct ggg gag cgg gcg tga 1722
 Asp Phe Ala Gln Ala Leu Ala Ser Ala Gly Glu Arg Ala
 565 570

<210> 36

<211> 573

<212> PRT

<213> *Deinococcus radiodurans*

<400> 36

Val Ala Phe Pro Cys Gly Gln Ala Gly Asn Lys Ile Thr Arg Pro Gly
 1 5 10 15

Gln Cys Val Asn Gly Gly Arg Ala Arg Ser Ala Pro Ser Lys Phe Arg
 20 25 30

Leu Asp Pro Leu Pro Ala Ser Ala Leu Leu Phe Glu Glu Leu Pro Gln
 35 40 45

Gln Glu Pro Pro Thr Ser Glu Arg Pro Arg Asp Leu Ala Arg Arg Arg
 50 55 60

Arg Gly Asn Arg Thr Pro Ser Arg His Gly Ala Lys Ala Ser Pro Leu
 65 70 75 80

Gly Val Ser Thr Met Ser Arg Arg Ala Gly Pro Ser Gly Lys Ala Thr
 85 90 95

Arg Met Thr Ile Phe Lys Cys Pro Phe Ser Ile His Ser Arg Gln Gly
 100 105 110

Glu Cys Arg Asp Trp Arg Pro Arg Thr Cys Phe Pro Arg Gly Ala Ala
 115 120 125

Thr Met Thr Asp Thr Lys Gln Pro Gln Pro Leu His Phe Glu Thr Leu
 130 135 140

Gln Val His Ala Gly Gln Arg Pro Asp Pro Val Thr Gly Ala Gln Gln
 145 150 155 160

Thr Pro Ile Tyr Ala Thr Asn Ser Tyr Val Phe Glu Ser Pro Glu His
165 170 175

Ala Ala Asp Leu Phe Gly Leu Arg Gln Phe Gly Asn Ile Tyr Ser Arg
180 185 190

Ile Met Asn Pro Thr Asn Asp Val Phe Glu Gln Arg Val Ala Ala Leu
195 200 205

Glu Gly Gly Val Gly Ala Leu Ser Val Ser Ser Gly His Ala Gly Gln
210 215 220

Leu Val Thr Leu Leu Thr Leu Ala Gln Ala Gly Asp Asn Ile Val Ser
225 230 235 240

Ser Pro Asn Leu Tyr Gly Gly Thr Val Asn Gln Phe Arg Val Thr Leu
245 250 255

Lys Arg Leu Gly Ile Glu Val Arg Phe Thr Ser Lys Asp Glu Arg Pro
260 265 270

Glu Glu Phe Ala Ala Leu Ile Asp Glu Arg Thr Arg Ala Val Tyr Leu
275 280 285

Glu Thr Ile Gly Asn Pro Ala Leu Asn Ile Pro Asp Phe Glu Gly Val
290 295 300

Ala Lys Val Ala His Glu His Gly Val Ala Val Val Val Asp Asn Thr
305 310 315 320

Phe Gly Ala Gly Gly Tyr Tyr Cys Gln Pro Leu Arg His Gly Ala Asn
325 330 335

Ile Val Leu His Ser Ala Ser Lys Trp Ile Gly Gly His Gly Asn Gly
340 345 350

Ile Gly Gly Val Ile Val Asp Gly Gly Asn Phe Asp Trp Gly Ser Gly
355 360 365

Arg Tyr Pro Leu Met Thr Glu Pro Ser Pro Ser Tyr His Gly Leu Lys
370 375 380

Phe Trp Glu Thr Phe Gly Glu Gly Asn Gly Leu Gly Leu Pro Asn Ile
385 390 395 400

Ala Phe Ile Thr Arg Ala Arg Thr Glu Gly Leu Arg Asp Leu Gly Thr
405 410 415

Thr Leu Ala Pro Gln Gln Ala Trp Gln Phe Leu Gln Gly Leu Glu Thr
420 425 430

Leu Ser Leu Arg Ala Glu Arg His Ala Glu Asn Thr Leu Ala Leu Ala
435 440 445

His Trp Leu Ile Ser His Pro Asp Val Lys Gln Val Thr Tyr Pro Gly
450 455 460

Leu Ser Asn His Pro His Tyr Asp Arg Ala Gln Thr Tyr Leu Pro Arg
465 470 475 480

Gly Ala Gly Ala Val Leu Thr Phe Glu Leu Arg Gly Gly Arg Ala Ala
485 490 495

Gly Glu Ala Phe Ile Arg Ser Val Lys Leu Ala Gln His Val Ala Asn
500 505 510

Val Gly Asp Thr Arg Thr Leu Val Ile His Pro Ala Ser Thr Thr His
515 520 525

Ser Gln Leu Asp Glu Val Thr Gln Thr Asn Ala Gly Val Thr Pro Gly
530 535 540

Leu Ile Arg Val Ser Val Gly Ile Glu His Val Asp Asp Ile Arg Glu
545 550 555 560

Asp Phe Ala Gln Ala Leu Ala Ser Ala Gly Glu Arg Ala
565 570

<210> 37

<211> 1284

<212> DNA

<213> Rhodobacter capsulatus

<220>

<221> CDS

<222> (1)..(1284)

<223>

<400> 37

atg acc gac cag gcc ttt gac acg ctg caa att cac gcg ggc gcc gaa Met Thr Asp Gln Ala Phe Asp Thr Leu Gln Ile His Ala Gly Ala Glu 1 5 10 15	48
ccc gat ccc gcg acg ggc gcg cgg cag gtg ccg att tac cag acc acc Pro Asp Pro Ala Thr Gly Ala Arg Gln Val Pro Ile Tyr Gln Thr Thr 20 25 30	96
tcc tat gtc ttc aag gac gcc gac cat gcc gcg cgc ctg ttc ggg ctg Ser Tyr Val Phe Lys Asp Ala Asp His Ala Ala Arg Leu Phe Gly Leu 35 40 45	144
cag gag gtg ggc tat atc tat tcc cgc ctg acc aac ccg acc gtt tcg Gln Glu Val Gly Tyr Ile Tyr Ser Arg Leu Thr Asn Pro Thr Val Ser 50 55 60	192
gca ctg gcc gcc cgc gtt gcg gcg ctt gaa ggc ggc gtg ggc gcg gtc Ala Leu Ala Ala Arg Val Ala Ala Leu Glu Gly Gly Val Gly Ala Val 65 70 75 80	240
tgc tgc tcg tcc ggc cat gcg gcg cag atc atg gcg ctg ttt ccg ctg Cys Cys Ser Ser Gly His Ala Ala Gln Ile Met Ala Leu Phe Pro Leu 85 90 95	288
atg ggg ccg ggg ctg aac atc gtc gcc tcg acc ccg ctt tac ggc ggc Met Gly Pro Gly Leu Asn Ile Val Ala Ser Thr Arg Leu Tyr Gly Gly 100 105 110	336
acg atc acc cag ttc agc cag acc atc aaa cgc ttc ggc tgg tcc tgc Thr Ile Thr Gln Phe Ser Gln Thr Ile Lys Arg Phe Gly Trp Ser Cys 115 120 125	384
acc ttt gtc gat ttc gac gat ctg gcg gcg ctc gag gcc gcg gtg gat Thr Phe Val Asp Phe Asp Asp Leu Ala Ala Leu Glu Ala Ala Val Asp 130 135 140	432
gac aac acc cgg gcg atc ttt tgc gaa tcg atc tcg aac ccg ggc ggc Asp Asn Thr Arg Ala Ile Phe Cys Glu Ser Ile Ser Asn Pro Gly Gly 145 150 155 160	480
tac atc acc gac ctg ccc gcc gtc gcg gcg gtg gcg aac aag gtc ggc Tyr Ile Thr Asp Leu Pro Ala Val Ala Ala Val Ala Asn Lys Val Gly 165 170 175	528
ctg ccg ctc att gtc gac aac acg ctg gcc tcg cct tat ctc tgc cgc Leu Pro Leu Ile Val Asp Asn Thr Leu Ala Ser Pro Tyr Leu Cys Arg 180 185 190	576
ccg atc gag cat ggc gcg acg ctg gtt gtc cat tcc gcc acg aaa tac Pro Ile Glu His Gly Ala Thr Leu Val Val His Ser Ala Thr Lys Tyr 195 200 205	624
ctg acc ggc aac ggc acg gtg acg ggc ggg gtg atc gtc gat tcg ggc Leu Thr Gly Asn Gly Thr Val Thr Gly Gly Val Ile Val Asp Ser Gly 210 215 220	672
aag ttc gac tgg tcg gcc tcg ggc aag ttc ccc agc ctt tcg gcg ccc Lys Phe Asp Trp Ser Ala Ser Gly Lys Phe Pro Ser Leu Ser Ala Pro 225 230 235 240	720
gaa ccc gcc tat cac ggg ctg aag ttc cac gag gca ctc ggc ccg atg Glu Pro Ala Tyr His Gly Leu Lys Phe His Glu Ala Leu Gly Pro Met 245 250 255	768

gcc ttc acc ttc cat tcg atc gcc gtc ggg ctg cgc gat ctg ggc atg 816
 Ala Phe Thr Phe His Ser Ile Ala Val Gly Leu Arg Asp Leu Gly Met
 260 265 270

acg atg aac ccg cag ggc gcg cat tac acg ctg atg ggg atc gag acg 864
 Thr Met Asn Pro Gln Gly Ala His Tyr Thr Leu Met Gly Ile Glu Thr
 275 280 285

ctc agc ctg cgc atg gac aag cac gtc gcc aat gcg aag gcg gtg gcg 912
 Leu Ser Leu Arg Met Asp Lys His Val Ala Asn Ala Lys Ala Val Ala
 290 295 300

gaa tgg ctg gcc aaa gac ccg cgc atc gac ttc gtc acc tgg gcc ggg 960
 Glu Trp Leu Ala Lys Asp Pro Arg Ile Asp Phe Val Thr Trp Ala Gly
 305 310 315 320

ctg ccc tcc tcg ccc tgg cac gaa cgc gcc gag cgg ctt tgc ccg aag 1008
 Leu Pro Ser Ser Pro Trp His Glu Arg Ala Glu Arg Leu Cys Pro Lys
 325 330 335

ggg gcg ggg gcg ctt ttc acc gtc gcg gtc aag ggc ggc tat gag gcc 1056
 Gly Ala Gly Ala Leu Phe Thr Val Ala Val Lys Gly Gly Tyr Glu Ala
 340 345 350

tgc gtg aaa ttg gtc aac aat ctc aag ctg ttc agc cat gtg gca aac 1104
 Cys Val Lys Leu Val Asn Asn Leu Lys Leu Phe Ser His Val Ala Asn
 355 360 365

ctg ggc gac gcg cgc tcg ctg atc atc cat tcg gcc tcg acc acg cac 1152
 Leu Gly Asp Ala Arg Ser Leu Ile Ile His Ser Ala Ser Thr Thr His
 370 375 380

cgt cag ctg acc gag gaa cag cag atc aag gcg ggg gcg gcg ccg aat 1200
 Arg Gln Leu Thr Glu Glu Gln Gln Ile Lys Ala Gly Ala Ala Pro Asn
 385 390 395 400

gtg gtg cgg ctc tcg atc ggg atc gag aat gcc gcc gat ctg atc gcc 1248
 Val Val Arg Leu Ser Ile Gly Ile Glu Asn Ala Ala Asp Leu Ile Ala
 405 410 415

gat ctg gat cag gct ctg gcc gcc gcc acc gcc tga 1284
 Asp Leu Asp Gln Ala Leu Ala Ala Ala Thr Ala
 420 425

<210> 38

<211> 427

<212> PRT

<213> Rhodobacter capsulatus

<400> 38

Met Thr Asp Gln Ala Phe Asp Thr Leu Gln Ile His Ala Gly Ala Glu
 1 5 10 15

Pro Asp Pro Ala Thr Gly Ala Arg Gln Val Pro Ile Tyr Gln Thr Thr

20

25

30

Ser Tyr Val Phe Lys Asp Ala Asp His Ala Ala Arg Leu Phe Gly Leu
35 40 45

Gln Glu Val Gly Tyr Ile Tyr Ser Arg Leu Thr Asn Pro Thr Val Ser
50 55 60

Ala Leu Ala Ala Arg Val Ala Ala Leu Glu Gly Gly Val Gly Ala Val
65 70 75 80

Cys Cys Ser Ser Gly His Ala Ala Gln Ile Met Ala Leu Phe Pro Leu
85 90 95

Met Gly Pro Gly Leu Asn Ile Val Ala Ser Thr Arg Leu Tyr Gly Gly
100 105 110

Thr Ile Thr Gln Phe Ser Gln Thr Ile Lys Arg Phe Gly Trp Ser Cys
115 120 125

Thr Phe Val Asp Phe Asp Asp Leu Ala Ala Leu Glu Ala Ala Val Asp
130 135 140

Asp Asn Thr Arg Ala Ile Phe Cys Glu Ser Ile Ser Asn Pro Gly Gly
145 150 155 160

Tyr Ile Thr Asp Leu Pro Ala Val Ala Ala Val Ala Asn Lys Val Gly
165 170 175

Leu Pro Leu Ile Val Asp Asn Thr Leu Ala Ser Pro Tyr Leu Cys Arg
180 185 190

Pro Ile Glu His Gly Ala Thr Leu Val Val His Ser Ala Thr Lys Tyr
195 200 205

Leu Thr Gly Asn Gly Thr Val Thr Gly Gly Val Ile Val Asp Ser Gly
210 215 220

Lys Phe Asp Trp Ser Ala Ser Gly Lys Phe Pro Ser Leu Ser Ala Pro
225 230 235 240

Glu Pro Ala Tyr His Gly Leu Lys Phe His Glu Ala Leu Gly Pro Met
245 250 255

Ala Phe Thr Phe His Ser Ile Ala Val Gly Leu Arg Asp Leu Gly Met
260 265 270

Thr Met Asn Pro Gln Gly Ala His Tyr Thr Leu Met Gly Ile Glu Thr
 275 280 285

Leu Ser Leu Arg Met Asp Lys His Val Ala Asn Ala Lys Ala Val Ala
 290 295 300

Glu Trp Leu Ala Lys Asp Pro Arg Ile Asp Phe Val Thr Trp Ala Gly
 305 310 315 320

Leu Pro Ser Ser Pro Trp His Glu Arg Ala Glu Arg Leu Cys Pro Lys
 325 330 335

Gly Ala Gly Ala Leu Phe Thr Val Ala Val Lys Gly Gly Tyr Glu Ala
 340 345 350

Cys Val Lys Leu Val Asn Asn Leu Lys Leu Phe Ser His Val Ala Asn
 355 360 365

Leu Gly Asp Ala Arg Ser Leu Ile Ile His Ser Ala Ser Thr Thr His
 370 375 380

Arg Gln Leu Thr Glu Glu Gln Gln Ile Lys Ala Gly Ala Ala Pro Asn
 385 390 395 400

Val Val Arg Leu Ser Ile Gly Ile Glu Asn Ala Ala Asp Leu Ile Ala
 405 410 415

Asp Leu Asp Gln Ala Leu Ala Ala Ala Thr Ala
 420 425

<210> 39

<211> 1269

<212> DNA

<213> Pasteurella multocida

<220>

<221> CDS

<222> (1) .. (1269)

<223>

<400> 39

atg gaa ttt gca aca aaa tgt cta cat gcc ggt tat aca ccg aaa aat
 Met Glu Phe Ala Thr Lys Cys Leu His Ala Gly Tyr Thr Pro Lys Asn

1	5	10	15	
ggt gag cct cgt gtt caa ccg atc gta caa agt acc act ttt acc tac				96
Gly Glu Pro Arg Val Gln Pro Ile Val Gln Ser Thr Thr Phe Thr Tyr				
20		25	30	
gat tcc gcc gaa gaa att ggt aag tta ttt gat tta caa gcg gct ggc				144
Asp Ser Ala Glu Glu Ile Gly Lys Leu Phe Asp Leu Gln Ala Ala Gly				
35		40	45	
tat ttt tac acc cgc ctt tca aat cct act acc aat gcg gca gaa gaa				192
Tyr Phe Tyr Thr Arg Leu Ser Asn Pro Thr Thr Asn Ala Ala Glu Glu				
50		55	60	
aaa att acc gca ctt gaa ggc ggt gta gca acc atg tgt acc gca tca				240
Lys Ile Thr Ala Leu Glu Gly Gly Val Ala Thr Met Cys Thr Ala Ser				
65		70	75	80
ggg caa gcc gcc gtg ttt tac gcg atg ctc aat att tta caa gcc ggt				288
Gly Gln Ala Ala Val Phe Tyr Ala Met Leu Asn Ile Leu Gln Ala Gly				
85		90	95	
gat cac ttt att tct tca tcg tat gtt tac ggt ggt agc tac aac tta				336
Asp His Phe Ile Ser Ser Ser Tyr Val Tyr Gly Gly Ser Tyr Asn Leu				
100		105	110	
ttt gca cat acc ttc aaa aaa atg gga att gag gtc act ttt gtg gat				384
Phe Ala His Thr Phe Lys Lys Met Gly Ile Glu Val Thr Phe Val Asp				
115		120	125	
caa gat tta cct ctt gag gaa tta aaa aaa gct att cgc cca aat acg				432
Gln Asp Leu Pro Leu Glu Glu Leu Lys Lys Ala Ile Arg Pro Asn Thr				
130		135	140	
aaa gcc att ttt gcc gaa act att gcc aat ccc gca tta cgc gtg ttg				480
Lys Ala Ile Phe Ala Glu Thr Ile Ala Asn Pro Ala Leu Arg Val Leu				
145		150	155	160
gat att gaa aag ttt gtt gca ctt gcg aag gca gca caa gcc cct tta				528
Asp Ile Glu Lys Phe Val Ala Leu Ala Lys Ala Ala Gln Ala Pro Leu				
165		170	175	
tta gtt gac aat act ttt gca acc ccg tat ttt tgt cgc cct atc gaa				576
Leu Val Asp Asn Thr Phe Ala Thr Pro Tyr Phe Cys Arg Pro Ile Glu				
180		185	190	
ttt ggt gct aac gtg gta att cat agt acg tca aaa tat tta gat ggg				624
Phe Gly Ala Asn Val Val Ile His Ser Thr Ser Lys Tyr Leu Asp Gly				
195		200	205	
cat gcg att gcg ttg gga ggt tcg atc aca gat ggc ggg aat ttt gat				672
His Ala Ile Ala Leu Gly Gly Ser Ile Thr Asp Gly Gly Asn Phe Asp				
210		215	220	
tgg aat aat ggt aaa ttc cca caa tta agc aca cct gat caa act tat				720
Trp Asn Asn Gly Lys Phe Pro Gln Leu Ser Thr Pro Asp Gln Thr Tyr				
225		230	235	240
cac ggt tta gtt tat acc gaa acc ttt gtt cca gcc gct tat att gtc				768
His Gly Leu Val Tyr Thr Glu Thr Phe Val Pro Ala Ala Tyr Ile Val				
245		250	255	
aaa gcc cgt gtg caa tta atg cgt gat tta ggt gcc aca cca gca cca				816

Lys Ala Arg Val Gln Leu Met Arg Asp Leu Gly Ala Thr Pro Ala Pro
 260 265 270
 caa aat agt ttc ttg ctc aat gtg ggc atg gaa act ctt gca ctg cgt 864
 Gln Asn Ser Phe Leu Leu Asn Val Gly Met Glu Thr Leu Ala Leu Arg
 275 280 285
 atg caa cgt cat tat gaa aat gca caa gcg gtc gcc gaa ttt tta gaa 912
 Met Gln Arg His Tyr Glu Asn Ala Gln Ala Val Ala Glu Phe Leu Glu
 290 295 300
 aat cat cca caa gtg gca aaa gtg agt tat ccg ggc ttg gca agt tca 960
 Asn His Pro Gln Val Ala Lys Val Ser Tyr Pro Gly Leu Ala Ser Ser
 305 310 315 320
 cct gat cat gca cta aaa caa aaa tat tta cca aac ggt tta tgt ggt 1008
 Pro Asp His Ala Leu Lys Gln Lys Tyr Leu Pro Asn Gly Leu Cys Gly
 325 330 335
 gtg att tcc ttt gaa att aga ggg gga aga gaa act gca gca aaa tgg 1056
 Val Ile Ser Phe Glu Ile Arg Gly Gly Arg Glu Thr Ala Ala Lys Trp
 340 345 350
 ctg aat gcg cta caa ctg gct tct cgt gaa gtc cat gta gcg gat att 1104
 Leu Asn Ala Leu Gln Leu Ala Ser Arg Glu Val His Val Ala Asp Ile
 355 360 365
 cgc act tgt gct tta cat ccg gcg acg tca aca cac cgt caa tta agt 1152
 Arg Thr Cys Ala Leu His Pro Ala Thr Ser Thr His Arg Gln Leu Ser
 370 375 380
 gag gct gaa tta gaa aaa gtg ggg att tct gcg ggt tta att cgt ctt 1200
 Glu Ala Glu Leu Glu Lys Val Gly Ile Ser Ala Gly Leu Ile Arg Leu
 385 390 395 400
 tct tgc ggt att gaa agt atc caa gat att ttg gct gac tta gaa caa 1248
 Ser Cys Gly Ile Glu Ser Ile Gln Asp Ile Leu Ala Asp Leu Glu Gln
 405 410 415
 gca ttc cac gcg gca aaa taa 1269
 Ala Phe His Ala Ala Lys
 420

<210> 40

<211> 422

<212> PRT

<213> Pasteurella multocida

<400> 40

Met Glu Phe Ala Thr Lys Cys Leu His Ala Gly Tyr Thr Pro Lys Asn
 1 5 10 15

Gly Glu Pro Arg Val Gln Pro Ile Val Gln Ser Thr Thr Phe Thr Tyr
 20 25 30

Asp Ser Ala Glu Glu Ile Gly Lys Leu Phe Asp Leu Gln Ala Ala Gly
35 40 45

Tyr Phe Tyr Thr Arg Leu Ser Asn Pro Thr Thr Asn Ala Ala Glu Glu
50 55 60

Lys Ile Thr Ala Leu Glu Gly Gly Val Ala Thr Met Cys Thr Ala Ser
65 70 75 80

Gly Gln Ala Ala Val Phe Tyr Ala Met Leu Asn Ile Leu Gln Ala Gly
85 90 95

Asp His Phe Ile Ser Ser Ser Tyr Val Tyr Gly Gly Ser Tyr Asn Leu
100 105 110

Phe Ala His Thr Phe Lys Lys Met Gly Ile Glu Val Thr Phe Val Asp
115 120 125

Gln Asp Leu Pro Leu Glu Glu Leu Lys Lys Ala Ile Arg Pro Asn Thr
130 135 140

Lys Ala Ile Phe Ala Glu Thr Ile Ala Asn Pro Ala Leu Arg Val Leu
145 150 155 160

Asp Ile Glu Lys Phe Val Ala Leu Ala Lys Ala Ala Gln Ala Pro Leu
165 170 175

Leu Val Asp Asn Thr Phe Ala Thr Pro Tyr Phe Cys Arg Pro Ile Glu
180 185 190

Phe Gly Ala Asn Val Val Ile His Ser Thr Ser Lys Tyr Leu Asp Gly
195 200 205

His Ala Ile Ala Leu Gly Gly Ser Ile Thr Asp Gly Gly Asn Phe Asp
210 215 220

Trp Asn Asn Gly Lys Phe Pro Gln Leu Ser Thr Pro Asp Gln Thr Tyr
225 230 235 240

His Gly Leu Val Tyr Thr Glu Thr Phe Val Pro Ala Ala Tyr Ile Val
245 250 255

Lys Ala Arg Val Gln Leu Met Arg Asp Leu Gly Ala Thr Pro Ala Pro
260 265 270

Gln Asn Ser Phe Leu Leu Asn Val Gly Met Glu Thr Leu Ala Leu Arg
275 280 285

Met Gln Arg His Tyr Glu Asn Ala Gln Ala Val Ala Glu Phe Leu Glu
 290 295 300

Asn His Pro Gln Val Ala Lys Val Ser Tyr Pro Gly Leu Ala Ser Ser
 305 310 315 320

Pro Asp His Ala Leu Lys Gln Lys Tyr Leu Pro Asn Gly Leu Cys Gly
 325 330 335

Val Ile Ser Phe Glu Ile Arg Gly Gly Arg Glu Thr Ala Ala Lys Trp
 340 345 350

Leu Asn Ala Leu Gln Leu Ala Ser Arg Glu Val His Val Ala Asp Ile
 355 360 365

Arg Thr Cys Ala Leu His Pro Ala Thr Ser Thr His Arg Gln Leu Ser
 370 375 380

Glu Ala Glu Leu Glu Lys Val Gly Ile Ser Ala Gly Leu Ile Arg Leu
 385 390 395 400

Ser Cys Gly Ile Glu Ser Ile Gln Asp Ile Leu Ala Asp Leu Glu Gln
 405 410 415

Ala Phe His Ala Ala Lys
 420

<210> 41

<211> 1266

<212> DNA

<213> Clostridium difficile

<220>

<221> CDS

<222> (1) .. (1266)

<223>

<400> 41

atg tat aat aaa gaa aca ata tgt gtg caa gga aat tat aaa cca ggt
 Met Tyr Asn Lys Glu Thr Ile Cys Val Gln Gly Asn Tyr Lys Pro Gly
 1 5 10 15

aat gga gaa cca aga gta cta cct tta tat caa agt aca act ttt aaa	96
Asn Gly Glu Pro Arg Val Leu Pro Leu Tyr Gln Ser Thr Thr Phe Lys	
20 25 30	
tat agc agt ata gac caa ctt gct gaa tta ttt gat tta aaa gtt gat	144
Tyr Ser Ser Ile Asp Gln Leu Ala Glu Leu Phe Asp Leu Lys Val Asp	
35 40 45	
gga cat ata tat tca aga ata agc aat cct act att caa gct ttt gaa	192
Gly His Ile Tyr Ser Arg Ile Ser Asn Pro Thr Ile Gln Ala Phe Glu	
50 55 60	
gaa aaa ata agt tta cta gag ggt gga gta tct tct gta gct gta tca	240
Glu Lys Ile Ser Leu Leu Glu Gly Gly Val Ser Ser Val Ala Val Ser	
65 70 75 80	
tca ggg caa tct gca aat atg ttg gca gtt tta aat ata tgt aaa tca	288
Ser Gly Gln Ser Ala Asn Met Leu Ala Val Leu Asn Ile Cys Lys Ser	
85 90 95	
gga gat agt ata ctt tgt tct tca aaa gta tat gga gga aca ttc aat	336
Gly Asp Ser Ile Leu Cys Ser Ser Lys Val Tyr Gly Gly Thr Phe Asn	
100 105 110	
tta cta gga cct agt ctt aaa aaa ttt ggt ata gat tta ata tcg ttt	384
Leu Leu Gly Pro Ser Leu Lys Lys Phe Gly Ile Asp Leu Ile Ser Phe	
115 120 125	
gac tta gat tca agt gaa gat gag ata gta gaa ctt gca aag gaa aat	432
Asp Leu Asp Ser Ser Glu Asp Glu Ile Val Glu Leu Ala Lys Glu Asn	
130 135 140	
act aag gtt gtg ttt gca gaa aca ctt gca aat cca act ctt gaa gtc	480
Thr Lys Val Val Phe Ala Glu Thr Leu Ala Asn Pro Thr Leu Glu Val	
145 150 155 160	
ata gat ttt gaa aaa ata gca aat gta gct aag aga att aat gtt cca	528
Ile Asp Phe Glu Lys Ile Ala Asn Val Ala Lys Arg Ile Asn Val Pro	
165 170 175	
ttt att gtt gat aat tca tta gca tct cca gtg ctt tgt aac cct tta	576
Phe Ile Val Asp Asn Ser Leu Ala Ser Pro Val Leu Cys Asn Pro Leu	
180 185 190	
aag tat gga gca aat ata gtt act cat tct acc aca aaa tat tta gat	624
Lys Tyr Gly Ala Asn Ile Val Thr His Ser Thr Thr Lys Tyr Leu Asp	
195 200 205	
ggg cat gct tca agt gtt gga gga att ata gtg gat ggt gga aac ttt	672
Gly His Ala Ser Ser Val Gly Gly Ile Ile Val Asp Gly Gly Asn Phe	
210 215 220	
aac tgg gat aat gga aaa ttt cca gaa tta gtt gag cca gac cca aca	720
Asn Trp Asp Asn Gly Lys Phe Pro Glu Leu Val Glu Pro Asp Pro Thr	
225 230 235 240	
tat cat ggt ata agc tat act caa aaa ttt gga aat gcc gca tat gca	768
Tyr His Gly Ile Ser Tyr Thr Gln Lys Phe Gly Asn Ala Ala Tyr Ala	
245 250 255	
act aaa gca aga gtt cag ttg ctt aga gac tat gga aat tgt tta agc	816
Thr Lys Ala Arg Val Gln Leu Leu Arg Asp Tyr Gly Asn Cys Leu Ser	
260 265 270	

cca ttc aat gcg tat ctt act aat tta aat gtt gaa aca cta cat ctt 864
 Pro Phe Asn Ala Tyr Leu Thr Asn Leu Asn Val Glu Thr Leu His Leu
 275 280 285

aga atg gag aga cat agt gaa aat gca ctt aaa ata gct aga ttt tta 912
 Arg Met Glu Arg His Ser Glu Asn Ala Leu Lys Ile Ala Arg Phe Leu
 290 295 300

gaa aaa cat gaa aat gta gat tgg att aat tac cca gga ctt gaa gat 960
 Glu Lys His Glu Asn Val Asp Trp Ile Asn Tyr Pro Gly Leu Glu Asp
 305 310 315 320

aac aag tat tat gag aat gcc aaa aag tat tta tca aga gga tgt agt 1008
 Asn Lys Tyr Tyr Glu Asn Ala Lys Lys Tyr Leu Ser Arg Gly Cys Ser
 325 330 335

ggg gtt tta tca ttt gga gta aga ggt ggg tta gaa aat gcc aaa aaa 1056
 Gly Val Leu Ser Phe Gly Val Arg Gly Gly Leu Glu Asn Ala Lys Lys
 340 345 350

ttt gtg gaa aaa tta cag ata gca tct ttg gtt aca cat gtt tca gat 1104
 Phe Val Glu Lys Leu Gln Ile Ala Ser Leu Val Thr His Val Ser Asp
 355 360 365

gta aga act tgt gtt ata cat cca gct tca act act cat aga caa tta 1152
 Val Arg Thr Cys Val Ile His Pro Ala Ser Thr Thr His Arg Gln Leu
 370 375 380

aca gaa gaa caa tta att gca tct gga gta ttg cct tca cta ata aga 1200
 Thr Glu Glu Gln Leu Ile Ala Ser Gly Val Leu Pro Ser Leu Ile Arg
 385 390 395 400

tta tct gtt gga ata gaa aat gta gag gat tta ata gct gat tta aat 1248
 Leu Ser Val Gly Ile Glu Asn Val Glu Asp Leu Ile Ala Asp Leu Asn
 405 410 415

caa gct tta aat ttc taa 1266
 Gln Ala Leu Asn Phe
 420

<210> 42

<211> 421

<212> PRT

<213> Clostridium difficile

<400> 42

Met Tyr Asn Lys Glu Thr Ile Cys Val Gln Gly Asn Tyr Lys Pro Gly
 1 5 10 15

Asn Gly Glu Pro Arg Val Leu Pro Leu Tyr Gln Ser Thr Thr Phe Lys
 20 25 30

Tyr Ser Ser Ile Asp Gln Leu Ala Glu Leu Phe Asp Leu Lys Val Asp

35

40

45

Gly His Ile Tyr Ser Arg Ile Ser Asn Pro Thr Ile Gln Ala Phe Glu
50 55 60

Glu Lys Ile Ser Leu Leu Glu Gly Gly Val Ser Ser Val Ala Val Ser
65 70 75 80

Ser Gly Gln Ser Ala Asn Met Leu Ala Val Leu Asn Ile Cys Lys Ser
85 90 95

Gly Asp Ser Ile Leu Cys Ser Ser Lys Val Tyr Gly Gly Thr Phe Asn
100 105 110

Leu Leu Gly Pro Ser Leu Lys Lys Phe Gly Ile Asp Leu Ile Ser Phe
115 120 125

Asp Leu Asp Ser Ser Glu Asp Glu Ile Val Glu Leu Ala Lys Glu Asn
130 135 140

Thr Lys Val Val Phe Ala Glu Thr Leu Ala Asn Pro Thr Leu Glu Val
145 150 155 160

Ile Asp Phe Glu Lys Ile Ala Asn Val Ala Lys Arg Ile Asn Val Pro
165 170 175

Phe Ile Val Asp Asn Ser Leu Ala Ser Pro Val Leu Cys Asn Pro Leu
180 185 190

Lys Tyr Gly Ala Asn Ile Val Thr His Ser Thr Thr Lys Tyr Leu Asp
195 200 205

Gly His Ala Ser Ser Val Gly Gly Ile Ile Val Asp Gly Gly Asn Phe
210 215 220

Asn Trp Asp Asn Gly Lys Phe Pro Glu Leu Val Glu Pro Asp Pro Thr
225 230 235 240

Tyr His Gly Ile Ser Tyr Thr Gln Lys Phe Gly Asn Ala Ala Tyr Ala
245 250 255

Thr Lys Ala Arg Val Gln Leu Leu Arg Asp Tyr Gly Asn Cys Leu Ser
260 265 270

Pro Phe Asn Ala Tyr Leu Thr Asn Leu Asn Val Glu Thr Leu His Leu
275 280 285

Arg Met Glu Arg His Ser Glu Asn Ala Leu Lys Ile Ala Arg Phe Leu
 290 295 300

Glu Lys His Glu Asn Val Asp Trp Ile Asn Tyr Pro Gly Leu Glu Asp
 305 310 315 320

Asn Lys Tyr Tyr Glu Asn Ala Lys Lys Tyr Leu Ser Arg Gly Cys Ser
 325 330 335

Gly Val Leu Ser Phe Gly Val Arg Gly Gly Leu Glu Asn Ala Lys Lys
 340 345 350

Phe Val Glu Lys Leu Gln Ile Ala Ser Leu Val Thr His Val Ser Asp
 355 360 365

Val Arg Thr Cys Val Ile His Pro Ala Ser Thr Thr His Arg Gln Leu
 370 375 380

Thr Glu Glu Gln Leu Ile Ala Ser Gly Val Leu Pro Ser Leu Ile Arg
 385 390 395 400

Leu Ser Val Gly Ile Glu Asn Val Glu Asp Leu Ile Ala Asp Leu Asn
 405 410 415

Gln Ala Leu Asn Phe
 420

<210> 43

<211> 1272

<212> DNA

<213> Campylobacter jejuni

<220>

<221> CDS

<222> (1) .. (1272)

<223>

<400> 43

atg aat ttc aat aaa gaa act tta gca tta cac gga gct tat aat ttt
 Met Asn Phe Asn Lys Glu Thr Leu Ala Leu His Gly Ala Tyr Asn Phe
 1 5 10 15

48

gat act caa aga agt att agt gtg cct ata tat caa aac act gcg tat
 Asp Thr Gln Arg Ser Ile Ser Val Pro Ile Tyr Gln Asn Thr Ala Tyr

96

20	25	30	
aat ttt gaa aat ttg gat caa gct gca gca agg ttt aat ctt caa gaa Asn Phe Glu Asn Leu Asp Gln Ala Ala Ala Arg Phe Asn Leu Gln Glu 35 40 45			144
ctt ggc aat att tac tca aga ctt agc aat cct aca agc gat gtt tta Leu Gly Asn Ile Tyr Ser Arg Leu Ser Asn Pro Thr Ser Asp Val Leu 50 55 60			192
gga caa aga ctt gct aat gtc gaa gga ggg gct ttt gga att cct gtt Gly Gln Arg Leu Ala Asn Val Glu Gly Gly Ala Phe Gly Ile Pro Val 65 70 75 80			240
gct agc ggt atg gca gct tgt ttt tat gct ctt atc aat tta gca agt Ala Ser Gly Met Ala Ala Cys Phe Tyr Ala Leu Ile Asn Leu Ala Ser 85 90 95			288
tcg gga gat aat gtc gcg tat tcg aac aaa att tat ggt ggg act caa Ser Gly Asp Asn Val Ala Tyr Ser Asn Lys Ile Tyr Gly Gly Thr Gln 100 105 110			336
act tta att tct cac aca ctt aaa aat ttt ggc ata gaa gct agg gaa Thr Leu Ile Ser His Thr Leu Lys Asn Phe Gly Ile Glu Ala Arg Glu 115 120 125			384
ttt gat atc gat gat tta gat agc ttg gaa aaa gtt ata gat caa aac Phe Asp Ile Asp Asp Leu Asp Ser Leu Glu Lys Val Ile Asp Gln Asn 130 135 140			432
aca aaa gcg att ttt ttc gaa agt ctt tca aat cct caa att gcc ata Thr Lys Ala Ile Phe Phe Glu Ser Leu Ser Asn Pro Gln Ile Ala Ile 145 150 155 160			480
gct gat ata gaa aaa ata aac caa ata gca aaa aaa cat aaa atc gtt Ala Asp Ile Glu Lys Ile Asn Gln Ile Ala Lys Lys His Lys Ile Val 165 170 175			528
agc att tgt gat aat acc gtt gct act cct ttc tta ctc caa cct ttt Ser Ile Cys Asp Asn Thr Val Ala Thr Pro Phe Leu Leu Gln Pro Phe 180 185 190			576
aaa cat ggc gtg gat gta atc gtg cat agt tta agt aaa tat gta agc Lys His Gly Val Asp Val Ile Val His Ser Leu Ser Lys Tyr Val Ser 195 200 205			624
ggt caa ggc act gct ttg ggt gga gca ctt ata gaa aga aaa gat tta Gly Gln Gly Thr Ala Leu Gly Gly Ala Leu Ile Glu Arg Lys Asp Leu 210 215 220			672
aac gac ttg ctt aaa aat aac gat aga tat aaa gct ttt aac act cct Asn Asp Leu Leu Lys Asn Asn Asp Arg Tyr Lys Ala Phe Asn Thr Pro 225 230 235 240			720
gat cca agt tat cat gga ctg aat tta aat aca ctt gat ttg ccg att Asp Pro Ser Tyr His Gly Leu Asn Leu Asn Thr Leu Asp Leu Pro Ile 245 250 255			768
ttt agt att aga gtc atc atc act tgg ctt aga gat cta gga gct agc Phe Ser Ile Arg Val Ile Ile Thr Trp Leu Arg Asp Leu Gly Ala Ser 260 265 270			816
tta gca cct caa aat gct tgg tta ctt tta caa gga ctt gaa acc ttg			864

Leu	Ala	Pro	Gln	Asn	Ala	Trp	Leu	Leu	Leu	Gln	Gly	Leu	Glu	Thr	Leu		
	275						280					285					
gca	gtg	cgt	ata	gaa	aaa	cac	agt	caa	aat	gct	gaa	aaa	gtt	gcg	aat	912	
Ala	Val	Arg	Ile	Glu	Lys	His	Ser	Gln	Asn	Ala	Glu	Lys	Val	Ala	Asn		
	290					295				300							
ttt	tta	aat	tct	cat	cct	gat	atc	aag	ggc	gta	aat	tat	cct	act	tta	960	
Phe	Leu	Asn	Ser	His	Pro	Asp	Ile	Lys	Gly	Val	Asn	Tyr	Pro	Thr	Leu		
	305				310				315						320		
gca	agt	aat	gct	tat	cat	aat	tta	ttt	aaa	aaa	tat	ttt	gat	aaa	aat	1008	
Ala	Ser	Asn	Ala	Tyr	His	Asn	Leu	Phe	Lys	Lys	Tyr	Phe	Asp	Lys	Asn		
			325						330					335			
ttt	gct	agc	ggg	ctt	tta	agc	ttt	gaa	gct	aaa	gat	tat	gag	cat	gct	1056	
Phe	Ala	Ser	Gly	Leu	Leu	Ser	Phe	Glu	Ala	Lys	Asp	Tyr	Glu	His	Ala		
			340					345					350				
aga	aga	att	tgt	gat	aaa	act	caa	ctt	ttc	tta	ctt	gct	gca	aat	ttg	1104	
Arg	Arg	Ile	Cys	Asp	Lys	Thr	Gln	Leu	Phe	Leu	Leu	Ala	Ala	Asn	Leu		
		355					360					365					
ggc	gat	agc	aag	tct	ttg	atc	atc	cat	cct	gct	tct	act	act	cat	tcg	1152	
Gly	Asp	Ser	Lys	Ser	Leu	Ile	Ile	His	Pro	Ala	Ser	Thr	Thr	His	Ser		
	370					375				380							
caa	cta	agc	gaa	gaa	gaa	ctc	caa	aaa	gca	ggc	att	acg	aaa	gct	act	1200	
Gln	Leu	Ser	Glu	Glu	Glu	Leu	Gln	Lys	Ala	Gly	Ile	Thr	Lys	Ala	Thr		
	385				390				395					400			
ata	cgc	tta	agc	ata	gga	ctt	gaa	aat	agc	gat	gat	ttg	ata	gcg	gat	1248	
Ile	Arg	Leu	Ser	Ile	Gly	Leu	Glu	Asn	Ser	Asp	Asp	Leu	Ile	Ala	Asp		
				405				410						415			
tta	aaa	caa	gct	ata	gaa	agt	taa									1272	
Leu	Lys	Gln	Ala	Ile	Glu	Ser											
			420														

<210> 44

<211> 423

<212> PRT

<213> Campylobacter jejuni

<400> 44

Met	Asn	Phe	Asn	Lys	Glu	Thr	Leu	Ala	Leu	His	Gly	Ala	Tyr	Asn	Phe
1			5						10					15	

Asp	Thr	Gln	Arg	Ser	Ile	Ser	Val	Pro	Ile	Tyr	Gln	Asn	Thr	Ala	Tyr
		20						25					30		

Asn	Phe	Glu	Asn	Leu	Asp	Gln	Ala	Ala	Ala	Arg	Phe	Asn	Leu	Gln	Glu
	35					40						45			

Leu Gly Asn Ile Tyr Ser Arg Leu Ser Asn Pro Thr Ser Asp Val Leu
50 55 60

Gly Gln Arg Leu Ala Asn Val Glu Gly Gly Ala Phe Gly Ile Pro Val
65 70 75 80

Ala Ser Gly Met Ala Ala Cys Phe Tyr Ala Leu Ile Asn Leu Ala Ser
85 90 95

Ser Gly Asp Asn Val Ala Tyr Ser Asn Lys Ile Tyr Gly Gly Thr Gln
100 105 110

Thr Leu Ile Ser His Thr Leu Lys Asn Phe Gly Ile Glu Ala Arg Glu
115 120 125

Phe Asp Ile Asp Asp Leu Asp Ser Leu Glu Lys Val Ile Asp Gln Asn
130 135 140

Thr Lys Ala Ile Phe Phe Glu Ser Leu Ser Asn Pro Gln Ile Ala Ile
145 150 155 160

Ala Asp Ile Glu Lys Ile Asn Gln Ile Ala Lys Lys His Lys Ile Val
165 170 175

Ser Ile Cys Asp Asn Thr Val Ala Thr Pro Phe Leu Leu Gln Pro Phe
180 185 190

Lys His Gly Val Asp Val Ile Val His Ser Leu Ser Lys Tyr Val Ser
195 200 205

Gly Gln Gly Thr Ala Leu Gly Gly Ala Leu Ile Glu Arg Lys Asp Leu
210 215 220

Asn Asp Leu Leu Lys Asn Asn Asp Arg Tyr Lys Ala Phe Asn Thr Pro
225 230 235 240

Asp Pro Ser Tyr His Gly Leu Asn Leu Asn Thr Leu Asp Leu Pro Ile
245 250 255

Phe Ser Ile Arg Val Ile Ile Thr Trp Leu Arg Asp Leu Gly Ala Ser
260 265 270

Leu Ala Pro Gln Asn Ala Trp Leu Leu Leu Gln Gly Leu Glu Thr Leu
275 280 285

Ala Val Arg Ile Glu Lys His Ser Gln Asn Ala Glu Lys Val Ala Asn
290 295 300

Phe Leu Asn Ser His Pro Asp Ile Lys Gly Val Asn Tyr Pro Thr Leu
 305 310 315 320

Ala Ser Asn Ala Tyr His Asn Leu Phe Lys Lys Tyr Phe Asp Lys Asn
 325 330 335

Phe Ala Ser Gly Leu Leu Ser Phe Glu Ala Lys Asp Tyr Glu His Ala
 340 345 350

Arg Arg Ile Cys Asp Lys Thr Gln Leu Phe Leu Leu Ala Ala Asn Leu
 355 360 365

Gly Asp Ser Lys Ser Leu Ile Ile His Pro Ala Ser Thr Thr His Ser
 370 375 380

Gln Leu Ser Glu Glu Glu Leu Gln Lys Ala Gly Ile Thr Lys Ala Thr
 385 390 395 400

Ile Arg Leu Ser Ile Gly Leu Glu Asn Ser Asp Asp Leu Ile Ala Asp
 405 410 415

Leu Lys Gln Ala Ile Glu Ser
 420

<210> 45

<211> 1041

<212> DNA

<213> Streptococcus pneumoniae

<220>

<221> CDS

<222> (1) .. (1041)

<223>

<400> 45

ttg agg aaa cca ggg aac att tat act cgt atc acc aat cct aca aca 48
 Leu Arg Lys Pro Gly Asn Ile Tyr Thr Arg Ile Thr Asn Pro Thr Thr
 1 5 10 15

gct gcc ctt gaa ggt ggt gtt gaa gcg cta gca aca gca tca ggt atg 96
 Ala Ala Leu Glu Gly Gly Val Glu Ala Leu Ala Thr Ala Ser Gly Met
 20 25 30

act gca gtg act tat acg att ttg gcg att gcc cat gct ggt gac cat Thr Ala Val Thr Tyr Thr Ile Leu Ala Ile Ala His Ala Gly Asp His 35 40 45	144
gta gtg gct gct tcg act att tac ggt gga acc ttc aat ctt ttg aaa Val Val Ala Ala Ser Thr Ile Tyr Gly Gly Thr Phe Asn Leu Leu Lys 50 55 60	192
gaa ccc ctt cct cgt tat ggt atc aca aca acc ttt ttc gat att gat Glu Pro Leu Pro Arg Tyr Gly Ile Thr Thr Thr Phe Phe Asp Ile Asp 65 70 75 80	240
aat ttg gag gaa gta gaa gca gct atc aaa gac aat acc aag ctt gtc Asn Leu Glu Glu Val Glu Ala Ala Ile Lys Asp Asn Thr Lys Leu Val 85 90 95	288
ttg att gaa acc ttg ggt aac ccc ttg att aat att cca gac ctg gaa Leu Ile Glu Thr Leu Gly Asn Pro Leu Ile Asn Ile Pro Asp Leu Glu 100 105 110	336
aaa ctg gca gag att gct cat aaa cat caa atc cca ctt gtg tca gac Lys Leu Ala Glu Ile Ala His Lys His Gln Ile Pro Leu Val Ser Asp 115 120 125	384
aat act ttt gca aca cct tat ttg att aac gtc ttc tct cat ggc gtt Asn Thr Phe Ala Thr Pro Tyr Leu Ile Asn Val Phe Ser His Gly Val 130 135 140	432
gac att gcc att cac tct gtg act aag ttt atc ggt ggg cat ggt aca Asp Ile Ala Ile His Ser Val Thr Lys Phe Ile Gly Gly His Gly Thr 145 150 155 160	480
act att gga gga ata att gtc gat agt ggt cgt ttt gac tgg acg gct Thr Ile Gly Gly Ile Ile Val Asp Ser Gly Arg Phe Asp Trp Thr Ala 165 170 175	528
tca ggg aaa ttc cct caa ttt gtt gac gag ggt cca agc tgc cac aat Ser Gly Lys Phe Pro Gln Phe Val Asp Glu Gly Pro Ser Cys His Asn 180 185 190	576
ttg agc tat act cgt gat gtg ggt gca gca gcc ttt att ata gct gtt Leu Ser Tyr Thr Arg Asp Val Gly Ala Ala Ala Phe Ile Ile Ala Val 195 200 205	624
cga gtt caa ttg ctt cgt gat aca ggt gca gcc ttg tca cca ttc aat Arg Val Gln Leu Leu Arg Asp Thr Gly Ala Ala Leu Ser Pro Phe Asn 210 215 220	672
gct ttc ctc ttg cta caa aga ctt gaa acc tct tca ctt cgt gtg gaa Ala Phe Leu Leu Leu Gln Arg Leu Glu Thr Ser Ser Leu Arg Val Glu 225 230 235 240	720
cgc cat gta caa aat gct gag aca att gtt gat ttt ctt gtc aac cat Arg His Val Gln Asn Ala Glu Thr Ile Val Asp Phe Leu Val Asn His 245 250 255	768
cct aag gta gaa aag gta aat tat cca aaa ctt gca gat agt cct tat Pro Lys Val Glu Lys Val Asn Tyr Pro Lys Leu Ala Asp Ser Pro Tyr 260 265 270	816
cat gcc ttg gct gag aaa tac ttg cca aaa ggt gtc ggt tca atc ttt His Ala Leu Ala Glu Lys Tyr Leu Pro Lys Gly Val Gly Ser Ile Phe 275 280 285	864

acc ttc cac gtc aaa ggt ggc gag gaa gaa gca cgc aag gtc att gat 912
 Thr Phe His Val Lys Gly Gly Glu Glu Glu Ala Arg Lys Val Ile Asp
 290 295 300

aat tta gaa atc ttt tct gac ctt gca aac gcg gca gat gct aaa tcg. 960
 Asn Leu Glu Ile Phe Ser Asp Leu Ala Asn Ala Ala Asp Ala Lys Ser
 305 310 315 320

ctt gtt gtc cat cca gca aca acc act cac ggt caa ttg tca gaa aaa 1008
 Leu Val Val His Pro Ala Thr Thr Thr His Gly Gln Leu Ser Glu Lys
 325 330 335

gac cta gaa gca gca ggt gtc aca cca aac taa 1041
 Asp Leu Glu Ala Ala Gly Val Thr Pro Asn
 340 345

<210> 46

<211> 346

<212> PRT

<213> Streptococcus pneumoniae

<400> 46

Leu Arg Lys Pro Gly Asn Ile Tyr Thr Arg Ile Thr Asn Pro Thr Thr
 1 5 10 15

Ala Ala Leu Glu Gly Gly Val Glu Ala Leu Ala Thr Ala Ser Gly Met
 20 25 30

Thr Ala Val Thr Tyr Thr Ile Leu Ala Ile Ala His Ala Gly Asp His
 35 40 45

Val Val Ala Ala Ser Thr Ile Tyr Gly Gly Thr Phe Asn Leu Leu Lys
 50 55 60

Glu Pro Leu Pro Arg Tyr Gly Ile Thr Thr Thr Phe Phe Asp Ile Asp
 65 70 75 80

Asn Leu Glu Glu Val Glu Ala Ala Ile Lys Asp Asn Thr Lys Leu Val
 85 90 95

Leu Ile Glu Thr Leu Gly Asn Pro Leu Ile Asn Ile Pro Asp Leu Glu
 100 105 110

Lys Leu Ala Glu Ile Ala His Lys His Gln Ile Pro Leu Val Ser Asp
 115 120 125

Asn Thr Phe Ala Thr Pro Tyr Leu Ile Asn Val Phe Ser His Gly Val

130

135

140

Asp Ile Ala Ile His Ser Val Thr Lys Phe Ile Gly Gly His Gly Thr
 145 150 155 160

Thr Ile Gly Gly Ile Ile Val Asp Ser Gly Arg Phe Asp Trp Thr Ala
 165 170 175

Ser Gly Lys Phe Pro Gln Phe Val Asp Glu Gly Pro Ser Cys His Asn
 180 185 190

Leu Ser Tyr Thr Arg Asp Val Gly Ala Ala Ala Phe Ile Ile Ala Val
 195 200 205

Arg Val Gln Leu Leu Arg Asp Thr Gly Ala Ala Leu Ser Pro Phe Asn
 210 215 220

Ala Phe Leu Leu Leu Gln Arg Leu Glu Thr Ser Ser Leu Arg Val Glu
 225 230 235 240

Arg His Val Gln Asn Ala Glu Thr Ile Val Asp Phe Leu Val Asn His
 245 250 255

Pro Lys Val Glu Lys Val Asn Tyr Pro Lys Leu Ala Asp Ser Pro Tyr
 260 265 270

His Ala Leu Ala Glu Lys Tyr Leu Pro Lys Gly Val Gly Ser Ile Phe
 275 280 285

Thr Phe His Val Lys Gly Gly Glu Glu Glu Ala Arg Lys Val Ile Asp
 290 295 300

Asn Leu Glu Ile Phe Ser Asp Leu Ala Asn Ala Ala Asp Ala Lys Ser
 305 310 315 320

Leu Val Val His Pro Ala Thr Thr Thr His Gly Gln Leu Ser Glu Lys
 325 330 335

Asp Leu Glu Ala Ala Gly Val Thr Pro Asn
 340 345

<210> 47

<211> 1335

<212> DNA

<213> Saccharomyces cerevisiae

<220>

<221> CDS

<222> (1)..(1335)

<223>

<400> 47

atg cca tct cat ttc gat act gtt caa cta cac gcc ggc caa gag aac	48
Met Pro Ser His Phe Asp Thr Val Gln Leu His Ala Gly Gln Glu Asn	
1 5 10 15	
cct ggt gac aat gct cac aga tcc aga gct gta cca att tac gcc acc	96
Pro Gly Asp Asn Ala His Arg Ser Arg Ala Val Pro Ile Tyr Ala Thr	
20 25 30	
act tct tat gtt ttc gaa aac tct aag cat ggt tcg caa ttg ttt ggt	144
Thr Ser Tyr Val Phe Glu Asn Ser Lys His Gly Ser Gln Leu Phe Gly	
35 40 45	
cta gaa gtt cca ggt tac gtc tat tcc cgt ttc caa aac cca acc agt	192
Leu Glu Val Pro Gly Tyr Val Tyr Ser Arg Phe Gln Asn Pro Thr Ser	
50 55 60	
aat gtt ttg gaa gaa aga att gct gct tta gaa ggt ggt gct gct gct	240
Asn Val Leu Glu Glu Arg Ile Ala Ala Leu Glu Gly Gly Ala Ala Ala	
65 70 75 80	
ttg gct gtt tcc tcc ggt caa gcc gct caa acc ctt gcc atc caa ggt	288
Leu Ala Val Ser Ser Gly Gln Ala Ala Gln Thr Leu Ala Ile Gln Gly	
85 90 95	
ttg gca cac act ggt gac aac atc gtt tcc act tct tac tta tac ggt	336
Leu Ala His Thr Gly Asp Asn Ile Val Ser Thr Ser Tyr Leu Tyr Gly	
100 105 110	
ggt act tat aac cag ttc aaa atc tcg ttc aaa aga ttt ggt atc gag	384
Gly Thr Tyr Asn Gln Phe Lys Ile Ser Phe Lys Arg Phe Gly Ile Glu	
115 120 125	
gct aga ttt gtt gaa ggt gac aat cca gaa gaa ttc gaa aag gtc ttt	432
Ala Arg Phe Val Glu Gly Asp Asn Pro Glu Glu Phe Glu Lys Val Phe	
130 135 140	
gat gaa aga acc aag gct gtt tat ttg gaa acc att ggt aat cca aag	480
Asp Glu Arg Thr Lys Ala Val Tyr Leu Glu Thr Ile Gly Asn Pro Lys	
145 150 155 160	
tac aat gtt ccg gat ttt gaa aaa att gtt gca att gct cac aaa cac	528
Tyr Asn Val Pro Asp Phe Glu Lys Ile Val Ala Ile Ala His Lys His	
165 170 175	
ggt att cca gtt gtc gtt gac aac aca ttt ggt gcc ggt ggt tac ttc	576
Gly Ile Pro Val Val Val Asp Asn Thr Phe Gly Ala Gly Gly Tyr Phe	
180 185 190	
tgt cag cca att aaa tac ggt gct gat att gta aca cat tct gct acc	624
Cys Gln Pro Ile Lys Tyr Gly Ala Asp Ile Val Thr His Ser Ala Thr	

195	200	205	
aaa tgg att ggt ggt cat ggt act act atc ggt ggt att att gtt gac Lys Trp Ile Gly Gly His Gly Thr Thr Ile Gly Gly Ile Ile Val Asp 210 215 220			672
tct ggt aag ttc cca tgg aag gac tac cca gaa aag ttc cct caa ttc Ser Gly Lys Phe Pro Trp Lys Asp Tyr Pro Glu Lys Phe Pro Gln Phe 225 230 235 240			720
tct caa cct gcc gaa gga tat cac ggt act atc tac aat gaa gcc tac Ser Gln Pro Ala Glu Gly Tyr His Gly Thr Ile Tyr Asn Glu Ala Tyr 245 250 255			768
ggt aac ttg gca tac atc gtt cat gtt aga act gaa cta tta aga gat Gly Asn Leu Ala Tyr Ile Val His Val Arg Thr Glu Leu Leu Arg Asp 260 265 270			816
ttg ggt cca ttg atg aac cca ttt gcc tct ttc ttg cta cta caa ggt Leu Gly Pro Leu Met Asn Pro Phe Ala Ser Phe Leu Leu Leu Gln Gly 275 280 285			864
gtt gaa aca tta tct ttg aga gct gaa aga cac ggt gaa aat gca ttg Val Glu Thr Leu Ser Leu Arg Ala Glu Arg His Gly Glu Asn Ala Leu 290 295 300			912
aag tta gcc aaa tgg tta gaa caa tcc cca tac gta tct tgg gtt tca Lys Leu Ala Lys Trp Leu Glu Gln Ser Pro Tyr Val Ser Trp Val Ser 305 310 315 320			960
tac cct ggt tta gca tct cat tct cat cat gaa aat gct aag aag tat Tyr Pro Gly Leu Ala Ser His Ser His His Glu Asn Ala Lys Lys Tyr 325 330 335			1008
cta tct aac ggt ttc ggt ggt gtc tta tct ttc ggt gta aaa gac tta Leu Ser Asn Gly Phe Gly Gly Val Leu Ser Phe Gly Val Lys Asp Leu 340 345 350			1056
cca aat gcc gac aag gaa act gac cca ttc aaa ctt tct ggt gct caa Pro Asn Ala Asp Lys Glu Thr Asp Pro Phe Lys Leu Ser Gly Ala Gln 355 360 365			1104
gtt gtt gac aat tta aag ctt gcc tct aac ttg gcc aat gtt ggt gat Val Val Asp Asn Leu Lys Leu Ala Ser Asn Leu Ala Asn Val Gly Asp 370 375 380			1152
gcc aag acc tta gtc att gct cca tac ttc act acc cac aaa caa tta Ala Lys Thr Leu Val Ile Ala Pro Tyr Phe Thr Thr His Lys Gln Leu 385 390 395 400			1200
aat gac aaa gaa aag ttg gca tct ggt gtt acc aag gac tta att cgt Asn Asp Lys Glu Lys Leu Ala Ser Gly Val Thr Lys Asp Leu Ile Arg 405 410 415			1248
gtc tct gtt ggt atc gaa ttt att gat gac att att gca gac ttc cag Val Ser Val Gly Ile Glu Phe Ile Asp Asp Ile Ile Ala Asp Phe Gln 420 425 430			1296
caa tct ttt gaa act gtt ttc gct ggc caa aaa cca tga Gln Ser Phe Glu Thr Val Phe Ala Gly Gln Lys Pro 435 440			1335

<210> 48

<211> 444

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 48

Met Pro Ser His Phe Asp Thr Val Gln Leu His Ala Gly Gln Glu Asn
1 5 10 15

Pro Gly Asp Asn Ala His Arg Ser Arg Ala Val Pro Ile Tyr Ala Thr
20 25 30

Thr Ser Tyr Val Phe Glu Asn Ser Lys His Gly Ser Gln Leu Phe Gly
35 40 45

Leu Glu Val Pro Gly Tyr Val Tyr Ser Arg Phe Gln Asn Pro Thr Ser
50 55 60

Asn Val Leu Glu Glu Arg Ile Ala Ala Leu Glu Gly Gly Ala Ala Ala
65 70 75 80

Leu Ala Val Ser Ser Gly Gln Ala Ala Gln Thr Leu Ala Ile Gln Gly
85 90 95

Leu Ala His Thr Gly Asp Asn Ile Val Ser Thr Ser Tyr Leu Tyr Gly
100 105 110

Gly Thr Tyr Asn Gln Phe Lys Ile Ser Phe Lys Arg Phe Gly Ile Glu
115 120 125

Ala Arg Phe Val Glu Gly Asp Asn Pro Glu Glu Phe Glu Lys Val Phe
130 135 140

Asp Glu Arg Thr Lys Ala Val Tyr Leu Glu Thr Ile Gly Asn Pro Lys
145 150 155 160

Tyr Asn Val Pro Asp Phe Glu Lys Ile Val Ala Ile Ala His Lys His
165 170 175

Gly Ile Pro Val Val Val Asp Asn Thr Phe Gly Ala Gly Gly Tyr Phe
180 185 190

Cys Gln Pro Ile Lys Tyr Gly Ala Asp Ile Val Thr His Ser Ala Thr
195 200 205

Lys Trp Ile Gly Gly His Gly Thr Thr Ile Gly Gly Ile Ile Val Asp
210 215 220

Ser Gly Lys Phe Pro Trp Lys Asp Tyr Pro Glu Lys Phe Pro Gln Phe
225 230 235 240

Ser Gln Pro Ala Glu Gly Tyr His Gly Thr Ile Tyr Asn Glu Ala Tyr
245 250 255

Gly Asn Leu Ala Tyr Ile Val His Val Arg Thr Glu Leu Leu Arg Asp
260 265 270

Leu Gly Pro Leu Met Asn Pro Phe Ala Ser Phe Leu Leu Leu Gln Gly
275 280 285

Val Glu Thr Leu Ser Leu Arg Ala Glu Arg His Gly Glu Asn Ala Leu
290 295 300

Lys Leu Ala Lys Trp Leu Glu Gln Ser Pro Tyr Val Ser Trp Val Ser
305 310 315 320

Tyr Pro Gly Leu Ala Ser His Ser His His Glu Asn Ala Lys Lys Tyr
325 330 335

Leu Ser Asn Gly Phe Gly Gly Val Leu Ser Phe Gly Val Lys Asp Leu
340 345 350

Pro Asn Ala Asp Lys Glu Thr Asp Pro Phe Lys Leu Ser Gly Ala Gln
355 360 365

Val Val Asp Asn Leu Lys Leu Ala Ser Asn Leu Ala Asn Val Gly Asp
370 375 380

Ala Lys Thr Leu Val Ile Ala Pro Tyr Phe Thr Thr His Lys Gln Leu
385 390 395 400

Asn Asp Lys Glu Lys Leu Ala Ser Gly Val Thr Lys Asp Leu Ile Arg
405 410 415

Val Ser Val Gly Ile Glu Phe Ile Asp Asp Ile Ile Ala Asp Phe Gln
420 425 430

Gln Ser Phe Glu Thr Val Phe Ala Gly Gln Lys Pro
435 440

<211> 1335

<212> DNA

<213> Kluyveromyces lactis

<220>

<221> CDS

<222> (1) .. (1335)

<223>

<400> 49

atg cca tct cac ttc gat act ttg caa ttg cac gct ggt caa gaa aag	48
Met Pro Ser His Phe Asp Thr Leu Gln Leu His Ala Gly Gln Glu Lys	
1 5 10 15	
act gct gat gct cat aac cca aga gcc gtc cca att tac gct acc act	96
Thr Ala Asp Ala His Asn Pro Arg Ala Val Pro Ile Tyr Ala Thr Thr	
20 25 30	
tct tac gtc ttc aac gac tct aag cat ggt gct caa ttg ttc ggt tta	144
Ser Tyr Val Phe Asn Asp Ser Lys His Gly Ala Gln Leu Phe Gly Leu	
35 40 45	
gaa act cca ggt tac att tac tct cgt att atg aac cct act cta gac	192
Glu Thr Pro Gly Tyr Ile Tyr Ser Arg Ile Met Asn Pro Thr Leu Asp	
50 55 60	
gtc ttg gaa aag aga ttg gca gcc tta gaa ggt ggt att gct gct ttg	240
Val Leu Glu Lys Arg Leu Ala Ala Leu Glu Gly Gly Ile Ala Ala Leu	
65 70 75 80	
gct act tct tct ggc caa gct gct caa acc ttg gct gtc act ggt ttg	288
Ala Thr Ser Ser Gly Gln Ala Ala Gln Thr Leu Ala Val Thr Gly Leu	
85 90 95	
gcc cac act ggt gac aat att gtc tct acc tct ttc tta tac ggt ggt	336
Ala His Thr Gly Asp Asn Ile Val Ser Thr Ser Phe Leu Tyr Gly Gly	
100 105 110	
act tat aac caa ttc aag gtt gcc ttc aag aga tta gga att gaa gct	384
Thr Tyr Asn Gln Phe Lys Val Ala Phe Lys Arg Leu Gly Ile Glu Ala	
115 120 125	
aga ttt gtc gat ggt gac aag cca gaa gac ttc gaa aag ttg ttc gat	432
Arg Phe Val Asp Gly Asp Lys Pro Glu Asp Phe Glu Lys Leu Phe Asp	
130 135 140	
gaa aag act aag gct ctc tat ctg gaa tct atc ggt aat cct aag tac	480
Glu Lys Thr Lys Ala Leu Tyr Leu Glu Ser Ile Gly Asn Pro Lys Tyr	
145 150 155 160	
aat gtc cca gac ttc gaa aag att gtt gct gtt gct cat aag cat ggt	528
Asn Val Pro Asp Phe Glu Lys Ile Val Ala Val Ala His Lys His Gly	
165 170 175	

atc cca gtt gtt gtt gac aac act ttc ggt gcc ggt ggt ttc ttc tgc Ile Pro Val Val Val Asp Asn Thr Phe Gly Ala Gly Gly Phe Phe Cys 180 185 190	576
caa cct atc aaa tac ggt gct gat atc gtt act cac tct gct acc aag Gln Pro Ile Lys Tyr Gly Ala Asp Ile Val Thr His Ser Ala Thr Lys 195 200 205	624
tgg atc ggt ggt cat ggt gtc acc gtt ggt ggt gtc atc att gac tct Trp Ile Gly Gly His Gly Val Thr Val Gly Gly Val Ile Ile Asp Ser 210 215 220	672
ggg aag ttc cca tgg aag gat tac ccg gaa aag ttc cct caa ttc tct Gly Lys Phe Pro Trp Lys Asp Tyr Pro Glu Lys Phe Pro Gln Phe Ser 225 230 235 240	720
cag cca tct gaa ggt tat cat ggt ttg atc ttc aat gat gcc ttt ggt Gln Pro Ser Glu Gly Tyr His Gly Leu Ile Phe Asn Asp Ala Phe Gly 245 250 255	768
cca gct gct ttc att ggt cat gta aga acc gaa ttg cta aga gat tta Pro Ala Ala Phe Ile Gly His Val Arg Thr Glu Leu Leu Arg Asp Leu 260 265 270	816
ggg cca gtg ttg agt cca ttc gct ggt ttc ttg ttg tta cag ggt ctt Gly Pro Val Leu Ser Pro Phe Ala Gly Phe Leu Leu Leu Gln Gly Leu 275 280 285	864
gaa act ttg tct cta aga ggt gaa aga cac ggt tcc aac gct ttg aag Glu Thr Leu Ser Leu Arg Gly Glu Arg His Gly Ser Asn Ala Leu Lys 290 295 300	912
ttg gct caa tac ttg gaa agt tct cca tac gtt tca tgg gtc tct tac Leu Ala Gln Tyr Leu Glu Ser Ser Pro Tyr Val Ser Trp Val Ser Tyr 305 310 315 320	960
cca ggt ttg cca tct cac tct cac cac gaa aac gct aag aaa tac ttg Pro Gly Leu Pro Ser His Ser His His Glu Asn Ala Lys Lys Tyr Leu 325 330 335	1008
gaa aat ggt ttc ggt ggt gtt tta tcc ttc ggt gtc aaa gat ttg cct Glu Asn Gly Phe Gly Gly Val Leu Ser Phe Gly Val Lys Asp Leu Pro 340 345 350	1056
aac gct tcc gag gaa tct gat cca ttc aag gct tct ggt gcc caa gtt Asn Ala Ser Glu Glu Ser Asp Pro Phe Lys Ala Ser Gly Ala Gln Val 355 360 365	1104
gtt gac aac ttg aag ctg gct tct aac ttg gca aac gtt ggt gac tcc Val Asp Asn Leu Lys Leu Ala Ser Asn Leu Ala Asn Val Gly Asp Ser 370 375 380	1152
aag acc ttg gtc att gct cca tac ttc act aca cat caa caa ttg acc Lys Thr Leu Val Ile Ala Pro Tyr Phe Thr Thr His Gln Gln Leu Thr 385 390 395 400	1200
gac gaa gaa aag tta gct tct ggt gtt acc aag gac ttg atc cgt gtt Asp Glu Glu Lys Leu Ala Ser Gly Val Thr Lys Asp Leu Ile Arg Val 405 410 415	1248
tct gtt ggt act gaa ttc att gac gac att att gct gac ttt gaa gca Ser Val Gly Thr Glu Phe Ile Asp Asp Ile Ile Ala Asp Phe Glu Ala 420 425 430	1296

tct ttc gct act gtc ttc aat ggc caa aaa cct gaa taa
Ser Phe Ala Thr Val Phe Asn Gly Gln Lys Pro Glu
435 440

1335

<210> 50

<211> 444

<212> PRT

<213> Kluyveromyces lactis

<400> 50

Met Pro Ser His Phe Asp Thr Leu Gln Leu His Ala Gly Gln Glu Lys
1 5 10 15

Thr Ala Asp Ala His Asn Pro Arg Ala Val Pro Ile Tyr Ala Thr Thr
20 25 30

Ser Tyr Val Phe Asn Asp Ser Lys His Gly Ala Gln Leu Phe Gly Leu
35 40 45

Glu Thr Pro Gly Tyr Ile Tyr Ser Arg Ile Met Asn Pro Thr Leu Asp
50 55 60

Val Leu Glu Lys Arg Leu Ala Ala Leu Glu Gly Gly Ile Ala Ala Leu
65 70 75 80

Ala Thr Ser Ser Gly Gln Ala Ala Gln Thr Leu Ala Val Thr Gly Leu
85 90 95

Ala His Thr Gly Asp Asn Ile Val Ser Thr Ser Phe Leu Tyr Gly Gly
100 105 110

Thr Tyr Asn Gln Phe Lys Val Ala Phe Lys Arg Leu Gly Ile Glu Ala
115 120 125

Arg Phe Val Asp Gly Asp Lys Pro Glu Asp Phe Glu Lys Leu Phe Asp
130 135 140

Glu Lys Thr Lys Ala Leu Tyr Leu Glu Ser Ile Gly Asn Pro Lys Tyr
145 150 155 160

Asn Val Pro Asp Phe Glu Lys Ile Val Ala Val Ala His Lys His Gly
165 170 175

Ile Pro Val Val Val Asp Asn Thr Phe Gly Ala Gly Gly Phe Phe Cys

180

185

190

Gln Pro Ile Lys Tyr Gly Ala Asp Ile Val Thr His Ser Ala Thr Lys
195 200 205

Trp Ile Gly Gly His Gly Val Thr Val Gly Gly Val Ile Ile Asp Ser
210 215 220

Gly Lys Phe Pro Trp Lys Asp Tyr Pro Glu Lys Phe Pro Gln Phe Ser
225 230 235 240

Gln Pro Ser Glu Gly Tyr His Gly Leu Ile Phe Asn Asp Ala Phe Gly
245 250 255

Pro Ala Ala Phe Ile Gly His Val Arg Thr Glu Leu Leu Arg Asp Leu
260 265 270

Gly Pro Val Leu Ser Pro Phe Ala Gly Phe Leu Leu Leu Gln Gly Leu
275 280 285

Glu Thr Leu Ser Leu Arg Gly Glu Arg His Gly Ser Asn Ala Leu Lys
290 295 300

Leu Ala Gln Tyr Leu Glu Ser Ser Pro Tyr Val Ser Trp Val Ser Tyr
305 310 315 320

Pro Gly Leu Pro Ser His Ser His His Glu Asn Ala Lys Lys Tyr Leu
325 330 335

Glu Asn Gly Phe Gly Gly Val Leu Ser Phe Gly Val Lys Asp Leu Pro
340 345 350

Asn Ala Ser Glu Glu Ser Asp Pro Phe Lys Ala Ser Gly Ala Gln Val
355 360 365

Val Asp Asn Leu Lys Leu Ala Ser Asn Leu Ala Asn Val Gly Asp Ser
370 375 380

Lys Thr Leu Val Ile Ala Pro Tyr Phe Thr Thr His Gln Gln Leu Thr
385 390 395 400

Asp Glu Glu Lys Leu Ala Ser Gly Val Thr Lys Asp Leu Ile Arg Val
405 410 415

Ser Val Gly Thr Glu Phe Ile Asp Asp Ile Ile Ala Asp Phe Glu Ala
420 425 430

Ser Phe Ala Thr Val Phe Asn Gly Gln Lys Pro Glu
 435 440

<210> 51

<211> 1323

<212> DNA

<213> Candida albicans

<220>

<221> CDS

<222> (1) .. (1323)

<223>

<400> 51

atg cct tct cac ttt gat aca ctt caa tta cat gct ggt caa cca gtt 48
 Met Pro Ser His Phe Asp Thr Leu Gln Leu His Ala Gly Gln Pro Val
 1 5 10 15

gaa aaa cca cac caa cca aga gcc cca cca att tat gca acc acc tcc 96
 Glu Lys Pro His Gln Pro Arg Ala Pro Pro Ile Tyr Ala Thr Thr Ser
 20 25 30

tat gtt ttc aat gac tct aaa cac ggt gct caa tta ttt ggt tta gaa 144
 Tyr Val Phe Asn Asp Ser Lys His Gly Ala Gln Leu Phe Gly Leu Glu
 35 40 45

acc cca gga tac att tac tcc aga att atg aat cca aca aac gat gtg 192
 Thr Pro Gly Tyr Ile Tyr Ser Arg Ile Met Asn Pro Thr Asn Asp Val
 50 55 60

ttt gaa caa aga att gct gcc ttg gaa ggt ggt att ggt gca ttg gcc 240
 Phe Glu Gln Arg Ile Ala Ala Leu Glu Gly Gly Ile Gly Ala Leu Ala
 65 70 75 80

act tct tct ggt caa tca gct caa ttc ttg gcc att gct ggg ttg gct 288
 Thr Ser Ser Gly Gln Ser Ala Gln Phe Leu Ala Ile Ala Gly Leu Ala
 85 90 95

cat gct ggt gat aac att atc agt aca tcc tac ttg tat ggt ggt act 336
 His Ala Gly Asp Asn Ile Ile Ser Thr Ser Tyr Leu Tyr Gly Gly Thr
 100 105 110

tat aat caa ttc aaa gtt gct ttc aaa cgt ttg ggc att gaa acc aaa 384
 Tyr Asn Gln Phe Lys Val Ala Phe Lys Arg Leu Gly Ile Glu Thr Lys
 115 120 125

ttc gtt aat ggt gac gcc gct gaa gat ttt gct aaa ttg att gac gac 432
 Phe Val Asn Gly Asp Ala Ala Glu Asp Phe Ala Lys Leu Ile Asp Asp
 130 135 140

aag aca aaa gct att tat att gaa acc att gga aac cct aaa tat aat 480
 Lys Thr Lys Ala Ile Tyr Ile Glu Thr Ile Gly Asn Pro Lys Tyr Asn

145	150	155	160	
gtt ccg gac ttt gaa aaa atc acc aaa ttg gcc cat gaa cac ggt att				528
Val Pro Asp Phe Glu Lys Ile Thr Lys Leu Ala His Glu His Gly Ile	165	170	175	
cct gtt gtt gtc gac aac act ttt ggt gct ggt gga ttt tta gtt aac				576
Pro Val Val Val Asp Asn Thr Phe Gly Ala Gly Gly Phe Leu Val Asn	180	185	190	
cca att gcc cac ggt gct gat att gtt gtt cat tct gct act aaa tgg				624
Pro Ile Ala His Gly Ala Asp Ile Val Val His Ser Ala Thr Lys Trp	195	200	205	
att ggt ggt cac ggt act aca att gct ggt gtt att gtt gat tcc ggt				672
Ile Gly Gly His Gly Thr Thr Ile Ala Gly Val Ile Val Asp Ser Gly	210	215	220	
aac ttc cca tgg acc gag tac cca gaa aaa tac cca caa ttc tct aaa				720
Asn Phe Pro Trp Thr Glu Tyr Pro Glu Lys Tyr Pro Gln Phe Ser Lys	225	230	235	240
cca tca gaa ggt tac cac ggg ttg atc ttg aat gat gct tta ggt aag				768
Pro Ser Glu Gly Tyr His Gly Leu Ile Leu Asn Asp Ala Leu Gly Lys	245	250	255	
gcc gca tac att ggt cac ttg aga att gaa ttg ttg aga gac ttg ggt				816
Ala Ala Tyr Ile Gly His Leu Arg Ile Glu Leu Leu Arg Asp Leu Gly	260	265	270	
cca gct ttg aat cca ttt gga agt ttt ttg ttg ttg caa ggt tta gaa				864
Pro Ala Leu Asn Pro Phe Gly Ser Phe Leu Leu Leu Gln Gly Leu Glu	275	280	285	
act ttg tct ttg aga gtt gaa aga caa tct gaa aat gct ttg aaa ttg				912
Thr Leu Ser Leu Arg Val Glu Arg Gln Ser Glu Asn Ala Leu Lys Leu	290	295	300	
gcc caa tgg ttg gaa aag aac cca aat gtt gag tct gtg tcc tat ttg				960
Ala Gln Trp Leu Glu Lys Asn Pro Asn Val Glu Ser Val Ser Tyr Leu	305	310	315	320
gga ttg cca tct cac gaa tcc cac gaa ttg agt aaa aaa tac ttg aac				1008
Gly Leu Pro Ser His Glu Ser His Glu Leu Ser Lys Lys Tyr Leu Asn	325	330	335	
aat gac gct aag tac ttt ggt ggt gct tta gca ttt act gtc aag gac				1056
Asn Asp Ala Lys Tyr Phe Gly Gly Ala Leu Ala Phe Thr Val Lys Asp	340	345	350	
atc acc aac acc tcc agc gac cca ttc aat gaa gcc tca cca aag ttg				1104
Ile Thr Asn Thr Ser Ser Asp Pro Phe Asn Glu Ala Ser Pro Lys Leu	355	360	365	
gtt gac aat ttg gag att gct tca aac ttg gct aat gtg ggt gac tct				1152
Val Asp Asn Leu Glu Ile Ala Ser Asn Leu Ala Asn Val Gly Asp Ser	370	375	380	
aag act ttg gtt att gct cca tgg ttt act aca cat caa caa ttg tct				1200
Lys Thr Leu Val Ile Ala Pro Trp Phe Thr Thr His Gln Gln Leu Ser	385	390	395	400
gat gaa gaa aag ttg gct tct ggt gtt acc aag ggc tta atc aga gtt				1248

Asp Glu Glu Lys Leu Ala Ser Gly Val Thr Lys Gly Leu Ile Arg Val
 405 410 415

tct act ggt act gaa tat att gat gat att att aac gac ttt gaa caa 1296
 Ser Thr Gly Thr Glu Tyr Ile Asp Asp Ile Ile Asn Asp Phe Glu Gln
 420 425 430

gca ttc aag aag gtt tat aac aac taa 1323
 Ala Phe Lys Lys Val Tyr Asn Asn
 435 440

<210> 52

<211> 440

<212> PRT

<213> Candida albicans

<400> 52

Met Pro Ser His Phe Asp Thr Leu Gln Leu His Ala Gly Gln Pro Val
 1 5 10 15

Glu Lys Pro His Gln Pro Arg Ala Pro Pro Ile Tyr Ala Thr Thr Ser
 20 25 30

Tyr Val Phe Asn Asp Ser Lys His Gly Ala Gln Leu Phe Gly Leu Glu
 35 40 45

Thr Pro Gly Tyr Ile Tyr Ser Arg Ile Met Asn Pro Thr Asn Asp Val
 50 55 60

Phe Glu Gln Arg Ile Ala Ala Leu Glu Gly Gly Ile Gly Ala Leu Ala
 65 70 75 80

Thr Ser Ser Gly Gln Ser Ala Gln Phe Leu Ala Ile Ala Gly Leu Ala
 85 90 95

His Ala Gly Asp Asn Ile Ile Ser Thr Ser Tyr Leu Tyr Gly Gly Thr
 100 105 110

Tyr Asn Gln Phe Lys Val Ala Phe Lys Arg Leu Gly Ile Glu Thr Lys
 115 120 125

Phe Val Asn Gly Asp Ala Ala Glu Asp Phe Ala Lys Leu Ile Asp Asp
 130 135 140

Lys Thr Lys Ala Ile Tyr Ile Glu Thr Ile Gly Asn Pro Lys Tyr Asn
 145 150 155 160

Val Pro Asp Phe Glu Lys Ile Thr Lys Leu Ala His Glu His Gly Ile
165 170 175

Pro Val Val Val Asp Asn Thr Phe Gly Ala Gly Gly Phe Leu Val Asn
180 185 190

Pro Ile Ala His Gly Ala Asp Ile Val Val His Ser Ala Thr Lys Trp
195 200 205

Ile Gly Gly His Gly Thr Thr Ile Ala Gly Val Ile Val Asp Ser Gly
210 215 220

Asn Phe Pro Trp Thr Glu Tyr Pro Glu Lys Tyr Pro Gln Phe Ser Lys
225 230 235 240

Pro Ser Glu Gly Tyr His Gly Leu Ile Leu Asn Asp Ala Leu Gly Lys
245 250 255

Ala Ala Tyr Ile Gly His Leu Arg Ile Glu Leu Leu Arg Asp Leu Gly
260 265 270

Pro Ala Leu Asn Pro Phe Gly Ser Phe Leu Leu Leu Gln Gly Leu Glu
275 280 285

Thr Leu Ser Leu Arg Val Glu Arg Gln Ser Glu Asn Ala Leu Lys Leu
290 295 300

Ala Gln Trp Leu Glu Lys Asn Pro Asn Val Glu Ser Val Ser Tyr Leu
305 310 315 320

Gly Leu Pro Ser His Glu Ser His Glu Leu Ser Lys Lys Tyr Leu Asn
325 330 335

Asn Asp Ala Lys Tyr Phe Gly Gly Ala Leu Ala Phe Thr Val Lys Asp
340 345 350

Ile Thr Asn Thr Ser Ser Asp Pro Phe Asn Glu Ala Ser Pro Lys Leu
355 360 365

Val Asp Asn Leu Glu Ile Ala Ser Asn Leu Ala Asn Val Gly Asp Ser
370 375 380

Lys Thr Leu Val Ile Ala Pro Trp Phe Thr Thr His Gln Gln Leu Ser
385 390 395 400

Asp Glu Glu Lys Leu Ala Ser Gly Val Thr Lys Gly Leu Ile Arg Val
405 410 415

Ser Thr Gly Thr Glu Tyr Ile Asp Asp Ile Ile Asn Asp Phe Glu Gln
 420 425 430

Ala Phe Lys Lys Val Tyr Asn Asn
 435 440

<210> 53

<211> 1290

<212> DNA

<213> Schizosaccharomyces pombe

<220>

<221> CDS

<222> (1) .. (1290)

<223>

<400> 53

atg cca gtc gag agt gaa cat ttc gaa act tta caa tta cat gct ggc 48
 Met Pro Val Glu Ser Glu His Phe Glu Thr Leu Gln Leu His Ala Gly
 1 5 10 15

caa gag cct gat gct gct acc agc tct cgt gcc gtt ccc atc tac gct 96
 Gln Glu Pro Asp Ala Ala Thr Ser Ser Arg Ala Val Pro Ile Tyr Ala
 20 25 30

act act tcc tat gtt ttc cgt gat tgc gac cat ggc ggc cgc ttg ttc 144
 Thr Thr Ser Tyr Val Phe Arg Asp Cys Asp His Gly Gly Arg Leu Phe
 35 40 45

gga tta cag gaa cca ggt tac atc tac tcg cgt atg atg aat ccc acc 192
 Gly Leu Gln Glu Pro Gly Tyr Ile Tyr Ser Arg Met Met Asn Pro Thr
 50 55 60

gcc gac gtt ttt gag aaa cgt att gcc gcc ttg gag cat ggc gct gct 240
 Ala Asp Val Phe Glu Lys Arg Ile Ala Ala Leu Glu His Gly Ala Ala
 65 70 75 80

gca atc gct act agt tcc ggt act tcc gct ctc ttc atg gct ttg acc 288
 Ala Ile Ala Thr Ser Ser Gly Thr Ser Ala Leu Phe Met Ala Leu Thr
 85 90 95

acg ttg gct aag gcc ggt gat aac att gtc tcc act tct tac ctt tat 336
 Thr Leu Ala Lys Ala Gly Asp Asn Ile Val Ser Thr Ser Tyr Leu Tyr
 100 105 110

ggg ggt act tac aac ctc ttc aag gtt acc ctg cct aga ttg gga att 384
 Gly Gly Thr Tyr Asn Leu Phe Lys Val Thr Leu Pro Arg Leu Gly Ile
 115 120 125

act acc aag ttt gtc aat ggt gat gat cct aat gat ctt gca gct cag Thr Thr Lys Phe Val Asn Gly Asp Asp Pro Asn Asp Leu Ala Ala Gln 130 135 140	432
att gat gaa aac aca aag gct gtt tac gtt gag tcc atc ggc aat ccc Ile Asp Glu Asn Thr Lys Ala Val Tyr Val Glu Ser Ile Gly Asn Pro 145 150 155 160	480
atg tac aac gtt ccc gat ttt gag cgt atc gct gag gtt gct cat gcc Met Tyr Asn Val Pro Asp Phe Glu Arg Ile Ala Glu Val Ala His Ala 165 170 175	528
gct ggt gtg cct tta atg gtc gat aac act ttt ggc ggc ggt ggt tat Ala Gly Val Pro Leu Met Val Asp Asn Thr Phe Gly Gly Gly Gly Tyr 180 185 190	576
ttg gtt cgt ccc att gac cac ggt gcc gat atc gtt acc cac tct gcc Leu Val Arg Pro Ile Asp His Gly Ala Asp Ile Val Thr His Ser Ala 195 200 205	624
act aag tgg atc ggt ggt cat ggc act act att ggc ggt gtg att gtt Thr Lys Trp Ile Gly Gly His Gly Thr Thr Ile Gly Gly Val Ile Val 210 215 220	672
gat agt ggt aag ttt gac tgg aag aag aac agc aag cgt ttc cct gaa Asp Ser Gly Lys Phe Asp Trp Lys Lys Asn Ser Lys Arg Phe Pro Glu 225 230 235 240	720
ttc aac gag cct cat ccc ggt tac cat ggc atg gtc ttt act gaa act Phe Asn Glu Pro His Pro Gly Tyr His Gly Met Val Phe Thr Glu Thr 245 250 255	768
ttt ggt aac ttg gca tat gct ttt gct tgc cgt act caa act ctc cgt Phe Gly Asn Leu Ala Tyr Ala Phe Ala Cys Arg Thr Gln Thr Leu Arg 260 265 270	816
gat gtt ggt ggc aat gcc aat cca ttc ggt gtc ttt ttg ctt ctt caa Asp Val Gly Gly Asn Ala Asn Pro Phe Gly Val Phe Leu Leu Leu Gln 275 280 285	864
ggt ctt gaa acg ctt tct ctt cgt atg gag cgt cac gtt caa aat gca Gly Leu Glu Thr Leu Ser Leu Arg Met Glu Arg His Val Gln Asn Ala 290 295 300	912
ttt gct ctt gca aaa tat ttg gaa aag cac ccc aag gtt aac tgg gtt Phe Ala Leu Ala Lys Tyr Leu Glu Lys His Pro Lys Val Asn Trp Val 305 310 315 320	960
tct tac cct ggt ctt gaa tct cac gtc tct cac aaa ctt gcc aag aag Ser Tyr Pro Gly Leu Glu Ser His Val Ser His Lys Leu Ala Lys Lys 325 330 335	1008
tac ttg aaa aat ggt tac ggc gcc gtt ctc agc ttt ggc gct aaa ggt Tyr Leu Lys Asn Gly Tyr Gly Ala Val Leu Ser Phe Gly Ala Lys Gly 340 345 350	1056
ggc cct gat caa agt cgt aag gta gtc aat gcc tta aag ctt gct agt Gly Pro Asp Gln Ser Arg Lys Val Val Asn Ala Leu Lys Leu Ala Ser 355 360 365	1104
cag ttg gcc aat gtt ggt gat gcc aaa act ttg gtt atc gct cct gcc Gln Leu Ala Asn Val Gly Asp Ala Lys Thr Leu Val Ile Ala Pro Ala 370 375 380	1152

tat acc act cat tta caa tta act gat gag gag caa att tct gcc ggt 1200
 Tyr Thr Thr His Leu Gln Leu Thr Asp Glu Glu Gln Ile Ser Ala Gly
 385 390 395 400

gtc act aag gat ctt att cgt gtg gcc gtc ggt att gag cac atc gat 1248
 Val Thr Lys Asp Leu Ile Arg Val Ala Val Gly Ile Glu His Ile Asp
 405 410 415

gat att atc gcc gac ttt gct caa gct ttg gaa gtt gcc taa 1290
 Asp Ile Ile Ala Asp Phe Ala Gln Ala Leu Glu Val Ala
 420 425

<210> 54

<211> 429

<212> PRT

<213> Schizosaccharomyces pombe

<400> 54

Met Pro Val Glu Ser Glu His Phe Glu Thr Leu Gln Leu His Ala Gly
 1 5 10 15

Gln Glu Pro Asp Ala Ala Thr Ser Ser Arg Ala Val Pro Ile Tyr Ala
 20 25 30

Thr Thr Ser Tyr Val Phe Arg Asp Cys Asp His Gly Gly Arg Leu Phe
 35 40 45

Gly Leu Gln Glu Pro Gly Tyr Ile Tyr Ser Arg Met Met Asn Pro Thr
 50 55 60

Ala Asp Val Phe Glu Lys Arg Ile Ala Ala Leu Glu His Gly Ala Ala
 65 70 75 80

Ala Ile Ala Thr Ser Ser Gly Thr Ser Ala Leu Phe Met Ala Leu Thr
 85 90 95

Thr Leu Ala Lys Ala Gly Asp Asn Ile Val Ser Thr Ser Tyr Leu Tyr
 100 105 110

Gly Gly Thr Tyr Asn Leu Phe Lys Val Thr Leu Pro Arg Leu Gly Ile
 115 120 125

Thr Thr Lys Phe Val Asn Gly Asp Asp Pro Asn Asp Leu Ala Ala Gln
 130 135 140

Ile Asp Glu Asn Thr Lys Ala Val Tyr Val Glu Ser Ile Gly Asn Pro

145 150 155 160

Met Tyr Asn Val Pro Asp Phe Glu Arg Ile Ala Glu Val Ala His Ala
 165 170 175

Ala Gly Val Pro Leu Met Val Asp Asn Thr Phe Gly Gly Gly Gly Tyr
 180 185 190

Leu Val Arg Pro Ile Asp His Gly Ala Asp Ile Val Thr His Ser Ala
 195 200 205

Thr Lys Trp Ile Gly Gly His Gly Thr Thr Ile Gly Gly Val Ile Val
 210 215 220

Asp Ser Gly Lys Phe Asp Trp Lys Lys Asn Ser Lys Arg Phe Pro Glu
225 230 235 240

Phe Asn Glu Pro His Pro Gly Tyr His Gly Met Val Phe Thr Glu Thr
 245 250 255

Phe Gly Asn Leu Ala Tyr Ala Phe Ala Cys Arg Thr Gln Thr Leu Arg
 260 265 270

Asp Val Gly Gly Asn Ala Asn Pro Phe Gly Val Phe Leu Leu Leu Gln
 275 280 285

Gly Leu Glu Thr Leu Ser Leu Arg Met Glu Arg His Val Gln Asn Ala
 290 295 300

Phe Ala Leu Ala Lys Tyr Leu Glu Lys His Pro Lys Val Asn Trp Val
305 310 315 320

Ser Tyr Pro Gly Leu Glu Ser His Val Ser His Lys Leu Ala Lys Lys
 325 330 335

Tyr Leu Lys Asn Gly Tyr Gly Ala Val Leu Ser Phe Gly Ala Lys Gly
 340 345 350

Gly Pro Asp Gln Ser Arg Lys Val Val Asn Ala Leu Lys Leu Ala Ser
 355 360 365

Gln Leu Ala Asn Val Gly Asp Ala Lys Thr Leu Val Ile Ala Pro Ala
 370 375 380

Tyr Thr Thr His Leu Gln Leu Thr Asp Glu Glu Gln Ile Ser Ala Gly
385 390 395 400

Val Thr Lys Asp Leu Ile Arg Val Ala Val Gly Ile Glu His Ile Asp
405 410 415

Asp Ile Ile Ala Asp Phe Ala Gln Ala Leu Glu Val Ala
420 425

<210> 55

<211> 52

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 55

cccgggatcc gctagcggcg cgccggccgg cccggtgtga aataccgcac ag 52

<210> 56

<211> 53

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 56

tctagactcg agcggccgcg gccggccttt aaattgaaga cgaaagggcc tcg 53

<210> 57

<211> 47

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 57

gagatctaga cccggggatc cgctagcggg ctgctaaagg aagcgga 47

<210> 58

<211> 38

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 58

gagaggcgcg ccgctagcgt gggcgaagaa ctccagca 38

<210> 59

<211> 34

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 59

gagagggcg cgcgcaaag tcccgcttcg tgaa

34

<210> 60

<211> 34

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 60

gagagggcg cgcgtcaagt cggccaagcc acgc

34

<210> 61

<211> 140

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 61

tcgaatttaa atctcgagag gcctgacgtc gggcccggtta ccacgcgtca tatgactagt 60
tcggacctag ggatcgcgtc gacatcgatg ctcttctgcg ttaattaaca attgggatcc 120
tctagaccgc ggatttaaatt 140

<210> 62

<211> 140

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 62

gatcatttaa atcccgggtc tagaggatcc caattgtaa ttaacgcaga agagcatcga 60
tgctgacgat atccctaggt ccgaactagt catatgacgc gtggtaccgc gcccgacgtc 120
aggcctctcg agatttaaatt 140

<210> 63

<211> 33

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 63

gagagcggcc gccgacctt tttaacccat cac

33

<210> 64

<211> 32

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 64

aggagcggcc gccatcggca ttttcttttg cg

32

<210> 65

<211> 5091

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Plasmid

<400> 65

gccgcgactg	ccttcgcgaa	gccttgcccc	gcggaaat	cctccaccga	gttcgtgcac	60
accctatgc	caagcttctt	tcaccctaaa	ttcgagagat	tggattctta	ccgtggaaat	120
tcttcgcaa	aatcgctccc	tgatcgccct	tgcgacgttg	gcgtcggtgc	cgctggttgc	180
gcttggttg	accgacttga	tcagcggccg	ctcgatttaa	atctcgagag	gcctgacgtc	240
gggcccggta	ccacgcgtca	tatgactagt	tcggacctag	ggatatcgtc	gacatcgatg	300
ctcttctgcg	tttaattaaca	attgggatcc	tctagaccgg	ggatttaa	cgctagcggg	360
ctgctaaagg	aagcgggaaca	cgtagaaagc	cagtcgcgag	aaacggtgct	gaccccggtat	420
gaatgtcagc	tactgggcta	tctggacaag	ggaaaacgca	agcgcgaaaga	gaaagcaggt	480
agcttgca	gggcttacat	ggcgatagct	agactggg	gttttatgga	cagcaagcga	540
accggaattg	ccagctgggg	cgccctctgg	taagggttggg	aagccctgca	aagtaaactg	600
gatggcttct	ttgcccga	ggatctgatg	gcgcagggga	tcaagatctg	atcaagagac	660
aggatgagga	tcgtttcgca	tgattgaaca	agatggattg	cacgcaggtt	ctccggccgc	720
ttgggtggag	aggctattcg	gctatgactg	ggcacaacag	acaatcggct	gctctgatgc	780
cgccgtgttc	cggtgtcag	cgcagggg	ccgggttctt	tttgtcaaga	ccgacctgtc	840
cggtgccctg	aatgaactgc	aggacgaggc	agcgcggcta	tcgtggctgg	ccacgacggg	900
cgcttccttg	gcagctgtgc	tcgacgttgt	caactgaagc	ggaaggga	ggctgctatt	960
gggcgaagtg	ccggggcagg	atctcctgtc	atctcacctt	gctcctgccg	agaaagtatc	1020
catcatggct	gatgcaatgc	ggcggtgca	tacgcttgat	ccggctacct	gccccattcga	1080
ccaccaagcg	aaacatcgca	tcgagcgagc	acgtactcgg	atggaagccg	gtcttgctcga	1140
tcaggatgat	ctggacgaag	agcatcaggg	cgtcgcgcca	gccgaactgt	tcgccaggtc	1200
caaggcgcgc	atgcccgacg	gcgaggatct	cgtcgtgacc	catggcgatg	cctgcttgcc	1260
gaatatcatg	gtggaaaatg	gccgcttttc	tggattcatt	gactgtggcc	ggctgggtgt	1320
ggcggaccgc	tatcaggaca	tagcgttggc	taccgctgat	attgctgaag	agcttggcgg	1380
cgaatgggct	gaccgcttcc	tcgtgcttta	cggtatcgcc	gctcccgatt	cgcagcgc	1440
cgccttctat	cgccttcttg	acgagttctt	ctgagcggga	ctctgggggt	cgaaatgacc	1500
gaccaagcga	cgcccaacct	gccatcacga	gatttcgatt	ccaccgccgc	cttctatgaa	1560
aggttgggct	tcggaatcgt	tttcggggac	gccggctgga	tgatcctcca	gcgcggggat	1620
ctcatgctgg	agttcttcgc	ccacgctagc	ggcgcgccgg	ccggcccggt	gtgaaatacc	1680
gcacagatgc	gtaaggagaa	aataccgcat	caggcgctct	tccgcttctt	cgctcactga	1740
ctcgctgcgc	tcggctcgttc	ggctgcggcg	agcggtatca	gctcactcaa	aggcggta	1800
acggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	1860
aaaggccagg	aaccgtaaaa	aggccgcgct	gctggcggtt	ttccataggc	tcggcccccc	1920
tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccgga	caggactata	1980
aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	2040
gcttaccgga	tacctgtccg	cctttctccc	ttcgggaagc	gtggcgcttt	ctcatagctc	2100
acgctgtagg	tatctcagtt	cggtgtaggt	cgttcgctcc	aagctgggct	gtgtgcaacc	2160
accccccgtt	cagcccagcc	gctgcgcctt	atccggtaac	tatcgtcttg	agtcacaacc	2220
ggtaagacac	gacttatcgc	cactggcagc	agccactggg	aacaggatta	gcagagcgag	2280
gtatgtaggc	ggtgctacag	agttcttgaa	gtgggtggct	aactacggct	acactagaag	2340
gacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	2400
ctcttgatcc	ggcaaacaaa	ccaccgctgg	tagcgggtgg	ttttttgttt	gcaagcagca	2460
gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cggggtctga	2520
cgctcagtg	aacgaaaact	cacgttaagg	gattttgggt	atgagattat	caaaaaggat	2580
cttcacctag	atccttttaa	aggccggccg	cgccgcgcga	aagtcccgc	tcgtgaaaat	2640
tttcgtgccc	cgtgattttc	cgccaaaaac	tttaacgaac	gttcgttata	atggtgtcat	2700
gaccttcacg	acgaagtact	aaaattggcc	cgaatcatca	gctatggatc	tctctgattg	2760
cgcgctggag	tccgacgcgc	tcgatgctgc	cgtcgaattta	aaaacgggtga	tcggattttt	2820
ccgagctctc	gatacgacgg	acgcgccagc	atcacgagac	tgggccagtg	ccgcgagcga	2880


```

cctagaaact ctcgtggcgg atcttgagga gctggctgac gagctgctg ctcggccagc 2940
gccaggagga cgcacagtag tggaggatgc aatcagttgc gcctactgcg gtggcctgat 3000
tcctccccgg cctgaccgcg gaggacggcg cgcaaaatat tgctcagatg cgtgtcgtgc 3060
cgcagccagc cgcgagcgcg ccaacaaacg ccacgccgag gagctggagg cggctaggtc 3120
gcaaatggcg ctggaagtgc gtcccccgag cgaaattttg gccatggtcg tcacagagct 3180
ggaagcggca gcgagaatta tcgcgatcgt ggcggtgccc gcaggcatga caaacatcgt 3240
aaatgcccgcg tttcgtgtgc cgtggccgccc caggacgtgt cagcgccgccc accacctgca 3300
ccgaatcggc agcagcgtcg cgcgtcgaaa aagcgcacag gcggcaagaa gcgataagct 3360
gcacgaatac ctgaaaaatg ttgaacgccc cgtgagcggt aactcacagg gcgtcggcta 3420
acccccagtc caaacctggg agaaagcgct caaaaatgac tctagcggat tcacgagaca 3480
ttgacacacc ggccctggaaa ttttcgctg atctgttcga caccatccc gagctcgcgc 3540
tgcgatcacg tggctggacg agcgaagacc gccgcgaatt cctcgctcac ctgggcagag 3600
aaaatttcca gggcagcaag acccgcgact tcgccagcgc ttggatcaaa gacccgga 3660
cggagaaaca cagccgaagt tataccgagt tggttcaaaa tcgcttgccc ggtgccagta 3720
tgttgctctg acgcacgcgc agcacgcagc cgtgcttgct ctggacattg atgtgccgag 3780
ccaccaggcc ggcggggaaaa tcgagcacgt aaaccccagag gtctacgcga ttttggagcg 3840
ctgggcacgc ctggaaaaag cgccagcttg gatcggcgtg aatccactga gcgggaaatg 3900
ccagctcatc tggtcatctg atccggtgta tgccgcagca ggcagtagca gcccgaaat 3960
gcgctgctg gctgcaacga ccgaggaaat gaccgcgtt ttccgctg accaggcttt 4020
ttcacatagg ctgagccgtg gccactgcac tctccgacga tccagccgt accgctggca 4080
tgccagcac aatcgctgg atcgcttagc tgatcttatg gaggttgctc gcatgatctc 4140
aggcacagaa aaacctaaaa aacgctatga gcaggagttt tctagcggac gggcacgtat 4200
cgaagcggca agaaaagcca ctgcggaagc aaaagcactt gccacgctg aagcaagcct 4260
gccgagcgcc gctgaagcgt ctggagagct gatcgacggc gtccgtgtcc tctggactgc 4320
tccagggcgt gccgcccgtg atgagacggc ttttcgccac gctttgactg tgggatacca 4380
gttaaaaagc gctggtgagc gcctaaaaa cccaagggt catcgagcct acgagcgtgc 4440
ctacaccgct gctcaggcgg tcggaggagg ccgtgagcct gatctgccg cggactgtga 4500
ccgcccagc gattggccgc gacgtgtgcg cggctacgtc gctaaaggcc agccagtcgt 4560
ccctgctcgt cagacagaga cgcagagcca gccgaggcga aaagctctgg ccactatggg 4620
aagacgtggc ggtaaaaagg ccgcagaacg ctggaaagac ccaaacagtg agtacgccc 4680
agcacagcga gaaaaactag ctaagtccag tcaacgacaa gctaggaaag ctaaaggaaa 4740
tcgcttgacc attgcagggt ggtttatgac tgttgaggga gagactggct cgtggccgac 4800
aatcaatgaa gctatgtctg aatttagcgt gtcacgtcag accgtgaata gagcacttaa 4860
ggctctgcggg cattgaactt ccacgaggag cccgtaagg ctctctcttg tgtgcatct 4920
cgtaggcaga aaacggttcc cccgtagggt cctctctctg gcctccttc taggtcggg 4980
tgattgctct tgaagctctc taggggggct cacaccatag gcagataacg ttccccaccg 5040
gctcgctcgt taagcgacac aggactgctc ccaaagatct tcaaagccac t 5091

```

<210> 66

<211> 4323

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Plasmid

<400> 66

```

tctctcagcg tatggttgct gcctgagctg tagttgcctt catcgatgaa ctgctgtaca 60
ttttgatagc tttttccgct accgtcaaa attgatttat aatcctctac accgttgatg 120
ttcaaagagc tgtctgatgc tgatacgta acttgtgcag ttgtcagtg ttgtttgccg 180
taatgtttac cggagaaatc agtgtagaat aaacggattt ttccgtcaga tgtaaatgtg 240
gctgaacctg accattcttg tgtttggtct tttaggatag aatcatttgc atcgaatttg 300
tcgctgtctt taaagacgcg gccagcgttt ttccagctgt caatagaagt ttcgccgact 360
ttttgataga acatgtaaat cgatgtgtca tccgcatttt taggatctcc ggctaatagca 420
aagacgatgt ggtagccgtg atagtttgcg acagtgccgt cagcgttttg taatggccag 480
ctgtcccaaa cgtccaggcc ttttgcagaa gagatatttt taatttgtga cgaatcaaat 540
tcagaaactt gatatttttc atttttttgc tgttcaggga tttgcagcat atcatggcgt 600
gtaatatggg aaatgccgta tgtttcctta tatggctttt ggttcgtttc tttcgcaaac 660
gcttgagttg cgctcctg cagcagtgcg gtagtaaagg ttaatactgt tgcttgtttt 720
gcaaaacttt tgatgttcac cgttcagtgc tcctttttta tgtactgtgt tagcggctctg 780
cttcttcacg ccctcctggt tgaagatggc aagttagtta cgcacaataa aaaaagacct 840

```

aaaatatgta aggggtgacg ccaaagtata cactttgccc tttacacatt ttaggtcttg 900
cctgctttat cagtaacaaa cccgcgcgat ttacttttcg acctcattct attagactct 960
cgtttggtt gcaactggtc tattttcctc ttttgtttga tagaaaatca taaaaggatt 1020
tgcagactac gggcctaaag aactaaaaaa tctatctgtt tcttttcatt ctctgtattt 1080
tttatagttt ctgttgcatg ggcataaagt tgctttttta atcacaattc agaaaatata 1140
ataatatctc atttactaa ataatagtga acggcaggta tatgtgatgg gttaaaaagg 1200
atcggcgcc gctcgattta aatctcgaga ggctgacgt cgggcccggg accacgcgtc 1260
atatgactag ttcggacctt gggatatcgt cgacatcgat gctcttctgc gttaattaac 1320
aattgggatc ctctagacc cggatttaaa tcgctagcgg gctgctaaag gaagcgggac 1380
acgtagaaa cagtcgcga gaaacgggtc tgaccccgga tgaatgtcag ctactgggct 1440
atctggacaa gggaaaacgc aagecgaaag agaaagcagg tagcttgacg tgggcttaca 1500
tggcgatagc tagactgggc ggttttatgg acagcaagcg aaccggaatt gccagctggg 1560
gcgccctctg gtaagggtgg gaagccctgc aaagtaaact ggatggcttt cttgccgcca 1620
aggatctgat ggcgcagggg atcaagatct gatcaagaga caggatgagg atcgtttcgc 1680
atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 1740
ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca 1800
gcgcaggggc gcccggttct ttttgctcaag accgacctgt ccggtgccct gaatgaactg 1860
caggacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg 1920
ctcgacgttg tcaactgaagc gggaaaggac tggctgctat tgggcgaagt gccggggcag 1980
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 2040
cggcggtcgc atacgttga tccggctacc tgccatttcg accaccaagc gaaacatcgc 2100
atcgagcag cagctactcg gatggaagcc ggtcttgctg atcaggatga tctggacgaa 2160
gagcatcagg ggtctcgccc agccgaactg ttcgccagcg tcaaggcgcg catgcccgac 2220
ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 2280
ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac 2340
atagcgttgg ctaccctgga tattgctgaa gagcttgccg gcgaatgggc tgaccgcttc 2400
ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt 2460
gacgagttct tctgagcggg actctggggg tcgaaatgac cgaccaagcg acgcccacc 2520
tgccatcacg agatttcgat tccaccgccg ccttctatga aaggttgggc ttcggaatcg 2580
ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gaggttcttcg 2640
cccacgctag cggcgccggc gccggccggc tgtgaaatac cgcacagatg cgtaaggaga 2700
aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 2760
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 2820
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 2880
aaggccgcgt tgctggcggt tttccatagg ctccgcccc ctgacgagca tcacaaaaat 2940
cgacgctcaa gtcagagggt gcgaaaccgg acaggactat aaagatacca ggcgtttccc 3000
cctggaagct ccctcgtagc ctctcctggt ccgaccctgc cgcttacggg atacctgtcc 3060
gcctttctcc ctctgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 3120
tcgggttagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac 3180
cgctgcgct tatccggtaa ctatcgctct gagtccaacc cggtaaagaca cgacttatcg 3240
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 3300
gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggatatctg 3360
gctctgctga agccagttac ctccggaaaa agagttggta gctcttgatc cggcaaaca 3420
accaccgctg gtagcgggtg tttttttggt tgcaagcagc agattacggc cagaaaaaaa 3480
ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg gaacgaaaac 3540
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 3600
aaggccggcc gcggccgcca tcggcatttt cttttgcgtt tttatttgtt aactgttaat 3660
tgtccttggt caaggatgct gtctttgaca acagatgttt tcttgccctt gatgttcagc 3720
aggaagctcg gcgcaaacgt tgattgtttg tctgcgtaga atcctctgtt tgtcatatag 3780
cttgtaatca cgacattggt tcctttcgct tgaggtacag cgaagtgtga gtaagtaaa 3840
gttacatcgt taggatcaag atccattttt aacacaaggc cagttttgtt cagcggcttg 3900
tatgggccag ttaaagaatt agaaacataa ccaagcatgt aaatatcgtt agacgtaatg 3960
ccgtcaatcg tcatttttga tccgcgggag tcagtgaaca ggtaccattt gccgttcatt 4020
ttaaagacgt tcgcgcgttc aatttcactt gttactgtgt tagatgcaat cagcggtttc 4080
atcactttt tcagtgtgta atcatcgttt agctcaatca taccgagagc gccgtttgct 4140
aactcagccg tgcgtttttt atcgtttttt agaagttttt gactttcttg acggaagaat 4200
gatgtgcttt tgccatagta tgctttgtta aataaagatt cttcgccctg gtagccatct 4260
tcagttccag tgtttgcttc aaatactaag tatttggtggc ctttatcttc tacgtagtga 4320
gga 4323

<210> 67
<211> 35
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 67
gagagagaga cgcgtcccag tggctgagac gcatc 35

<210> 68
<211> 34
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 68
ctctctctgt cgacgaattc aatcttacgg cctg 34

<210> 69
<211> 5860
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:Plasmid

<400> 69
cccgggtacca cgcgtcccag tggctgagac gcatccgcta aagccccagg aaccctgtgc 60
agaaagaaaa cactcctctg gctaggtaga cacagtttat aaaggtagag ttgagcgggt 120
aactgtcagc acgtagatcg aaaggtgcac aaaggtggcc ctggtcgtac agaaatatgg 180
cggttcctcg cttgagagtg cggaacgcat tagaaacgtc gctgaacgga tcgttgccac 240
caagaaggct ggaaatgatg tcgtggttgt ctgctccgca atgggagaca ccacggatga 300
acttctagaa cttgcagcgg cagtgaatcc cgttccgcca gctcgtgaaa tggatatgct 360
cctgactgct ggtgagcgta tttctaacgc tctcgtcgcc atggctattg agtcccttgg 420
cgcagaagcc caatctttca cgggctctca ggctgggtgtg ctcaccaccg agcgcacagg 480
aaacgcacgc attgttgatg tcactccagg tcgtgtgcgt gaagcactcg atgagggcaa 540
gatctgcatt gttgctgggt tccagggtgt taataaagaa acccgcgatg tcaccacgtt 600
gggtcgtgggt ggttctgaca ccactgcagt tgcgttgga gctgctttga acgctgatgt 660
gtgtgagatt tactcggacg ttgacggtgt gtataccgct gaccgcgca tcgttcctaa 720
tgcacagaag ctggaaaagc tcagcttcga agaaatgctg gaacttgctg ctggttgctc 780
caagattttg gtgctgcgca gtgttgaata cgctcgtgca ttcaatgtgc cacttcgcgt 840
acgctcgtct tatagtaatg atcccggcac tttgattgcc ggctctatgg aggatattcc 900

tgtggaagaa gcagtcctta ccggtgtcgc aaccgacaag tccgaagcca aagtaaccgt 960
tctgggtatt tccgataagc caggcgaggc tgcgaaggtt ttccgtgcgt tggctgatgc 1020
agaaatcaac attgacatgg ttctgcagaa cgtctcttct gtagaagacg gcaccaccga 1080
catcaccttc acctgccctc gtcccgacgg ccgccgcgcg atggagatct tgaagaagct 1140
tcaggttcag ggcaactgga ccaatgtgct ttacgacgac caggtcggca aagtctccct 1200
cgtgggtgct ggcatgaagt ctacccagg tgttaccgca gattcatgg aagctctgcg 1260
cgatgtcaac gtgaacatcg aattgatttc cacctctgag attcgtatct ccgtgctgat 1320
ccgtgaagat gatctggatg ctgctgcacg tgcattgcat gacgattcc agctgggcgg 1380
cgaagacgaa gccgtcgttt atgcaggcac cggacgctaa agttttaaag gagtagtttt 1440
acaatgacca ccatcgcatg tgttggtgca accggccagg tcggccagggt tatgcgaccc 1500
cttttggaag agcgcaattt cccagctgac actgttcgtt tctttgcttc cccacgttcc 1560
gcaggccgta agattgaatt cgtcgacatc gatgctcttc tgcgttaatt aacaattggg 1620
atcctctaga cccgggattt aaatcgctag cgggctgcta aaggaagcgg aacacgtaga 1680
aagccagtcc gcagaaacgg tgctgacccc ggatgaatgt cagctactgg gctatctgga 1740
caagggaaaa cgcaagcgca aagagaaagc aggtagcttg cagtgggctt acatggcgat 1800
agctagactg ggcggtttta tggacagcaa gcgaaccgga attgccagct ggggcgcctt 1860
ctggttaagg tgggaagccc tgcaaagtaa actggatggc tttcttgccg ccaaggatct 1920
gatggcgtag gggatcaaga tctgatcaag agacaggatg aggatcgctt cgcattgatg 1980
aacaagatgg attgcacgca ggttctccgg ccgcttgggt ggagaggcta ttcggctatg 2040
actgggcaca acagacaatc ggctgctctg atgccgcgt gttccggctg tcagcgcagg 2100
ggcgcccggt tctttttgtc aagaccgacc tgtccggtgc cctgaatgaa ctgcaggacg 2160
aggcagcgcg gctatcgctg ctggccacga cgggcgttcc ttgcgcagct gtgctcgacg 2220
ttgtcactga agcgggaagg gactggctgc tattgggcga agtgccgggg caggatctcc 2280
tgtcatctca ccttgctcct gccgagaaag tatccatcat ggctgatgca atgcggcggc 2340
tgcatacgtt tgatccgggt acctgcccac tcgaccacca agcgaaacat cgcattcgagc 2400
gagcacgtac tcggatggaa gccgggtctg tcgatcagga tgatctggac gaagagcatc 2460
aggggctcgc gccagccgaa ctgttcgcca ggctcaaggc gcgcattgcc gacggcgagg 2520
atctcgtcgt gacctatggc gatgcctgct tgccgaatat catggtggaa aatggccgct 2580
tttctggatt catcgactgt ggccggctgg gtgtggcgga ccgctatcag gacatagcgt 2640
tggctacccg tgatattgct gaagagcttg gcggcgaaat ggctgaccgc ttcctcgtgc 2700
tttacgggat cgcgcgtccc gattcgcagc gcacgcctt ctatgcctt cttgacgagt 2760
tcttctgagc gggactctgg ggttcgaaat gaccgaccaa gcgacgccc acctgcccac 2820

acgagatttc	gattccaccg	ccgccttcta	tgaaagggtg	ggcttcggaa	tcgttttccg	2880
ggacgcgggc	tggatgatcc	tccagcgcg	ggatctcatg	ctggagttct	tcgcccacgc	2940
tagcggcgcg	ccggccggcc	cgggtgtgaa	taccgcacag	atgcgtaagg	agaaaatacc	3000
gcatcaggcg	ctcttccgct	tcctcgctca	ctgactcgct	gcgctcggtc	gttcggctgc	3060
ggcgagcggg	atcagctcac	tcaaaggcgg	taatacgggt	atccacagaa	tcaggggata	3120
acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	3180
cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatcacaaa	aatcgacgct	3240
caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcgttt	cccctggaa	3300
gctccctcgt	gcgctctcct	gttccgacce	tgccgcttac	cggatacctg	tcgcctttc	3360
tcccttcggg	aagcgtggcg	ctttctcata	gctcacgctg	taggtatctc	agttcgggtg	3420
aggctcgttc	ctccaagctg	ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	3480
ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	acacgactta	tcgccactgg	3540
cagcagccac	tggtaacagg	attagcagag	cgaggtatgt	aggcggtgct	acagagttct	3600
tgaagtgggt	gcctaactac	ggctacacta	gaaggacagt	atctgggtatc	tcgcgtctgc	3660
tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	3720
ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	gcgcagaaaa	aaaggatctc	3780
aagaagatcc	tttgatcttt	tctacggggg	ctgacgctca	gtggaacgaa	aactcacggt	3840
aagggtttt	ggtcatgaga	ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaaggccg	3900
gccgcggccg	ccatcggcct	tttcttttgc	gttttttatt	gttaactggt	aattgtcctt	3960
gttcaaggat	gctgtctttg	acaacagatg	ttttcttgcc	tttgatgttc	agcaggaagc	4020
tcggcgcaaa	cgttgattgt	ttgtctgcgt	agaatcctct	gtttgtcata	tagcttgtaa	4080
tcacgacatt	gtttcctttc	gcttgaggta	cagcgaagtg	tgagtaagta	aaggttacat	4140
cgttaggata	aagatccatt	tttaacacaa	ggccagtttt	gttcagcggc	ttgtatgggc	4200
cagttaaaga	attagaaaca	taaccaagca	tgtaaataatc	gttagacgta	atgccgtcaa	4260
tcgtcatttt	tgatccgcgg	gagtcagtga	acaggtaacca	tttgccgttc	attttaaaga	4320
cgttcgcgcg	ttcaatttca	tctgttactg	tggttagatgc	aatcagcggg	ttcatcactt	4380
ttttcagtg	gtaatcatcg	tttagctcaa	tcataccgag	agcgccggtt	gctaactcag	4440
ccgtgcggtt	tttatcgctt	tgccagaagt	tttgactttc	ttgacggaag	aatgatgtgc	4500
ttttgccata	gtatgctttg	ttaaataaag	attcttcgcc	ttggtagcca	tcttcagttc	4560
cagtgtttgc	ttcaaatact	aagtattttg	ggcctttatc	ttctacgtag	tgaggatctc	4620
tcagcgtatg	gttgctgcct	gagctgtagt	tgcccttcac	gatgaactgc	tgtacatttt	4680

gatacgtttt tccgtcacccg tcaaagattg atttataatc ctctacaccg ttgatgttca 4740
aagagctgtc tgatgctgat acgttaactt gtgcagttgt cagtgtttgt ttgccgtaat 4800
gtttaccgga gaaatcagtg tagaataaac ggatttttcc gtcagatgta aatgtggctg 4860
aacctgacca ttcttgtgtt tgggtctttta ggatagaatc atttgcacgc aatttgcgc 4920
tgtcttttaa gacgcggcca gcgtttttcc agctgtcaat agaagtttcg ccgacttttt 4980
gatagaacat gtaaactgat gtgtcatccg catttttagg atctccggt aatgcaaaga 5040
cgatgtggta gccgtgatag tttgcgacag tgccgtcagc gttttgtaat ggccagctgt 5100
cccaaacgtc caggcctttt gcagaagaga tatttttaaat tgtggacgaa tcaaattcag 5160
aaacttgata tttttcattt ttttgcgtgt cagggatttg cagcatatca tggcgtgtaa 5220
tatgggaaat gccgtatgtt tccttatatg gcttttggtt cgtttcttcc gcaaacgctt 5280
gagttgcgcc tcctgccagc agtgcggtag taaagggttaa tactgttgct tgttttgcaa 5340
actttttgat gttcatcgtt catgtctcct ttttatgta ctgtgttagc ggtctgcttc 5400
ttccagccct cctgtttgaa gatggcaagt tagttacgca caataaaaaa agacctaaaa 5460
tatgtaagggt gtgacgcaa agtatacact ttgcccttta cacatttttag gtcttgccctg 5520
ctttatcagt aacaaacccg cgcgatttac ttttcgacct cattctatta gactctcgtt 5580
tggattgcaa ctgggtctatt ttctctttt gtttgataga aaatcataaa aggatttgca 5640
gactacgggc ctaaagaact aaaaaatcta tctgtttctt ttcattctct gtatttttta 5700
tagtttctgt tgcattgggca taaagttgcc tttttaatca caattcagaa aatatcataa 5760
tatctcattt cactaaataa tagtgaacgg caggtatatg tgatgggtta aaaaggatcg 5820
gcggccgctc gatttaaate tcgagaggcc tgacgtcggg 5860

<210> 70

<211> 38

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 70

cggcaccacc gacatcatct tcacctgccc tcgttccg

38

<210> 71

<211> 38

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 71

cggaacgagg gcaggtgaag atgatgtcgg tgggtgccg

38

<210> 72
 <211> 1266
 <212> DNA
 <213> LysC Mutante
 <220>
 <221> CDS
 <222> (1) .. (1266)
 <223>

<400> 72

gtg gcc ctg gtc gta cag aaa tat ggc ggt tcc tcg ctt gag agt gcg 48
 Val Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala
 1 5 10 15

gaa cgc att aga aac gtc gct gaa cgg atc gtt gcc acc aag aag gct 96
 Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala
 20 25 30

gga aat gat gtc gtg gtt gtc tgc tcc gca atg gga gac acc acg gat 144
 Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp
 35 40 45

gaa ctt cta gaa ctt gca gcg gca gtg aat ccc gtt ccg cca gct cgt 192
 Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg
 50 55 60

gaa atg gat atg ctc ctg act gct ggt gag cgt att tct aac gct ctc 240
 Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu
 65 70 75 80

gtc gcc atg gct att gag tcc ctt ggc gca gaa gcc caa tct ttc acg 288
 Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr
 85 90 95

ggc tct cag gct ggt gtg ctc acc acc gag cgc cac gga aac gca cgc 336
 Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg
 100 105 110

att gtt gat gtc act cca ggt cgt gtg cgt gaa gca ctc gat gag ggc 384
 Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly
 115 120 125

aag atc tgc att gtt gct ggt ttc cag ggt gtt aat aaa gaa acc cgc 432
 Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg
 130 135 140

gat gtc acc acg ttg ggt cgt ggt ggt tct gac acc act gca gtt gcg 480
 Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala
 145 150 155 160

ttg gca gct gct ttg aac gct gat gtg tgt gag att tac tcg gac gtt 528
 Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val
 165 170 175

gac ggt gtg tat acc gct gac ccg cgc atc gtt cct aat gca cag aag 576
 Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys
 180 185 190

ctg gaa aag ctc agc ttc gaa gaa atg ctg gaa ctt gct gct gtt ggc 624
 Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly
 195 200 205

tcc aag att ttg gtg ctg cgc agt gtt gaa tac gct cgt gca ttc aat	672
Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn	
210 215 220	
gtg cca ctt cgc gta cgc tcg tct tat agt aat gat ccc ggc act ttg	720
Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu	
225 230 235 240	
att gcc ggc tct atg gag gat att cct gtg gaa gaa gca gtc ctt acc	768
Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr	
245 250 255	
ggg gtc gca acc gac aag tcc gaa gcc aaa gta acc gtt ctg ggt att	816
Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile	
260 265 270	
tcc gat aag cca ggc gag gct gcg aag gtt ttc cgt gcg ttg gct gat	864
Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp	
275 280 285	
gca gaa atc aac att gac atg gtt ctg cag aac gtc tct tct gta gaa	912
Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu	
290 295 300	
gac ggc acc acc gac atc atc ttc acc tgc cct cgt tcc gac ggc cgc	960
Asp Gly Thr Thr Asp Ile Ile Phe Thr Cys Pro Arg Ser Asp Gly Arg	
305 310 315 320	
cgc gcg atg gag atc ttg aag aag ctt cag gtt cag ggc aac tgg acc	1008
Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr	
325 330 335	
aat gtg ctt tac gac gac cag gtc ggc aaa gtc tcc ctc gtg ggt gct	1056
Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala	
340 345 350	
ggc atg aag tct cac cca ggt gtt acc gca gag ttc atg gaa gct ctg	1104
Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu	
355 360 365	
cgc gat gtc aac gtg aac atc gaa ttg att tcc acc tct gag att cgt	1152
Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg	
370 375 380	
att tcc gtg ctg atc cgt gaa gat gat ctg gat gct gct gca cgt gca	1200
Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Arg Ala	
385 390 395 400	
ttg cat gag cag ttc cag ctg ggc ggc gaa gac gaa gcc gtc gtt tat	1248
Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr	
405 410 415	
gca ggc acc gga cgc taa	1266
Ala Gly Thr Gly Arg	
420	

<210> 73

<211> 421

<212> PRT

<213> LysC Mutante

<400> 73

Val Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala
1 5 10 15

Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala
20 25 30

Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp
35 40 45

Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg
50 55 60

Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu
65 70 75 80

Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr
85 90 95

Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg
100 105 110

Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly
115 120 125

Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg
130 135 140

Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala
145 150 155 160

Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val
165 170 175

Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys
180 185 190

Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly
195 200 205

Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn
210 215 220

Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu
225 230 235 240

Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr

245

250

255

Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile
260 265 270

Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp
275 280 285

Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu
290 295 300

Asp Gly Thr Thr Asp Ile Ile Phe Thr Cys Pro Arg Ser Asp Gly Arg
305 310 315 320

Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr
325 330 335

Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala
340 345 350

Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu
355 360 365

Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg
370 375 380

Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala
385 390 395 400

Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr
405 410 415

Ala Gly Thr Gly Arg
420

<210> 74

<211> 5860

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Plasmid

<400> 74

cccgggtacca cgcgtcccgag tggtgagac gcatccgcta aagccccagg aaccctgtgc 60
agaaagaaaa cactcctctg gctaggtaga cacagtttat aaaggtagag ttgagcgggt 120
aactgtcagc acgtagatcg aaaggtgcac aaaggtggcc ctggtcgtac agaaatatgg 180
cggttcctcg cttgagagtg cggaacgcat tagaaacgtc gctgaacgga tcgttgccac 240

caagaaggct	ggaaatgatg	tcgtggttgt	ctgctccgca	atgggagaca	ccacggatga	300
acttctagaa	cttgacgcgg	cagtgaatcc	cggtccgcca	gctcgtgaaa	tggatatgct	360
cctgactgct	ggtgagcgta	tttctaacgc	tctcgtcgcc	atggctattg	agtcaccttg	420
cgcagaagcc	caatctttca	cgggctctca	ggctgggtgtg	ctcaccaccg	agcgccacgg	480
aaacgcacgc	attggtgatg	tactccagg	tcgtgtgctg	gaagcactcg	atgagggcaa	540
gatctgcatt	gttgctggtt	tccaggggtg	taataaagaa	acccgcgatg	tcaccacgtt	600
gggtcgtggt	ggttctgaca	ccactgcagt	tgcgttgcca	gctgctttga	acgctgatgt	660
gtgtgagatt	tactcggacg	ttgacgggtg	gtataccgct	gacccgcgca	tcgttcctaa	720
tgcacagaag	ctggaaaagc	tcagcttcga	agaaatgctg	gaacttgctg	ctgttggttc	780
caagattttg	gtgctgcgca	gtgttgaata	cgctcgtgca	ttcaatgtgc	cacttcgcgt	840
acgctcgtct	tatagtaatg	atcccgccac	tttgattgcc	ggctctatgg	aggatattcc	900
tgtggaagaa	gcagtcctta	cgggtgtcgc	aaccgacaag	tccgaagcca	aagtaaccgt	960
tctgggtatt	tccgataagc	caggcgaggc	tgcgaagggt	ttccgtgcgt	tggctgatgc	1020
agaaatcaac	attgacatgg	ttctgcagaa	cgctctctct	gtagaagacg	gcaccaccga	1080
catcatcttc	acctgccctc	gttcgcagcg	ccgccgcgcg	atggagatct	tgaagaagct	1140
tcagggttcag	ggcaactgga	ccaatgtgct	ttacgacgac	caggtcggca	aagtctccct	1200
cgtgggtgct	ggcatgaagt	ctcaccacag	tgttaccgca	gagttcatgg	aagctctgcg	1260
cgatgtcaac	gtgaacatcg	aattgatttc	cacctctgag	attcgtatct	ccgtgctgat	1320
ccgtgaagat	gatctggatg	ctgctgcacg	tgcattgcat	gagcagttcc	agctgggcgg	1380
cgaagacgaa	gccgtcgttt	atgcaggcac	cggacgctaa	agttttaaag	gagtagtttt	1440
acaatgacca	ccatcgcagt	tgttggtgca	accggccagg	tcggccaggt	tatgcgcacc	1500
cttttggaag	agcgcaatct	cccagctgac	actgttcggt	tctttgcttc	cccacgttcc	1560
gcaggccgta	agattgaatt	cgctgcacat	gatgctcttc	tgcgttaatt	aacaattggg	1620
atcctctaga	cccgggattt	aaatcgctag	cgggctgcta	aaggaagcgg	aacacgtaga	1680
aagccagtcc	gcagaaacgg	tgctgacccc	ggatgaatgt	cagctactgg	gctatctgga	1740
caagggaaaa	cgcaagcgca	aagagaaagc	aggtagcttg	cagtgggctt	acatggcgat	1800
agctagactg	ggcggtttta	tggacagcaa	gcgaaccgga	attgccagct	ggggcgccct	1860
ctggtgaagg	tgggaagccc	tgcaaagtaa	actggatggc	tttcttgccg	ccaaggatct	1920
gatggcgcat	gggatcaaga	tctgatcaag	agacaggatg	aggatcggtt	cgcgatgattg	1980
aacaagatgg	attgcacgca	ggttctccgg	ccgcttgggg	ggagaggcta	ttcggctatg	2040
actgggcaca	acagacaatc	ggctgctctg	atgccgccgt	gttcgggctg	tcagcgcagg	2100

ggcgccccgt tctttttgtc aagaccgacc tgtccggtgc cctgaatgaa ctgcaggacg 2160
aggcagcgcg gctatcgtgg ctggccacga cgggcgttcc ttgcgcagct gtgctcgacg 2220
ttgtcactga agcgggaagg gactggctgc tattgggcga agtgccgggg caggatctcc 2280
tgtcatctca ccttgctcct gccgagaaag tatccatcat ggctgatgca atgcggcgge 2340
tgcatacgct tgatccggct acctgcccat tcgaccacca agcgaaacat cgcatacgagc 2400
gagcacgtac tcggatggaa gccggtcttg tcgatcagga tgatctggac gaagagcatc 2460
aggggctcgc gccagccgaa ctgttcgcca ggctcaaggc gcgcatgccc gacggcgagg 2520
atctcgtcgt gacccatggc gatgcctgct tgccgaatat catggtggaa aatggccgct 2580
tttctggatt catcgactgt ggccggctgg gtgtggcgga ccgctatcag gacatagcgt 2640
tggetacccg tgatattgct gaagagcttg gcggcgaatg ggctgaccgc ttectcgtgc 2700
tttacggtat cgccgctccc gattcgcage gcatagcctt ctatcgctt cttgacgagt 2760
tcttctgagc gggactctgg ggttcgaaat gaccgaccaa gcgacgcca acctgccatc 2820
acgagatttc gattccaccg ccgccttcta tgaaaggctg ggcttcggaa tcgttttccg 2880
ggacgccggc tggatgatcc tccagcgcg ggatctcatg ctggagttct tcgcccacgc 2940
tagcggcgcg ccggccggcc cgggtgtgaaa taccgcacag atgcgtaagg agaaaatacc 3000
gcatcaggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 3060
ggcgagcggg atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 3120
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 3180
cgttgctggc gtttttccat aggtccgcgc cccctgacga gcatcacaaa aatcgacgct 3240
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa 3300
gctccctcgt gcgctctcct gttccgacce tgccgcttac cggatacctg tccgccttcc 3360
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg 3420
aggctgctcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 3480
ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg 3540
cagcagccac tggtaacagg attagcagag cgaggatatgt aggcgggtgct acagagttct 3600
tgaagtgggt gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 3660
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 3720
ctggtagcgg tgggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 3780
aagaagatcc tttgatcttt tctacggggg ctgacgctca gtggaacgaa aactcacggt 3840
aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaaggccg 3900
gccgcggccg ccatacgcat tttcttttgc gtttttattt gtttaactgtt aattgtcctt 3960
gttcaaggat gctgtctttg acaacagatg tttctttgcc tttgatgttc agcaggaagc 4020

tcggcgcaaa cggttgattgt ttgtctgctt agaatectct gtttgtcata tagcttgtaa 4080
tcacgacatt gtttcctttc gcttgaggta cagcgaagtg tgagtaagta aaggttacat 4140
cgtaggatac aagatccatt tttaacacaa ggccagtttt gttcagcggc ttgtatgggc 4200
cagttaaaga attagaaaca taaccaagca tgtaaataac gttagacgta atgccgtcaa 4260
tcgtcatttt tgatccgogg gagtcagtga acaggtagca tttgccgttc attttaaaga 4320
cgttcgcgog ttcaattttca tctgttactg tgtagatgc aatcagcggg ttcactactt 4380
ttttcagtgt gtaatcatcg tttagctcaa tcataccgag agcgccggtt gctaactcag 4440
ccgtgcggtt tttatcgctt tgcagaagtt tttgactttc ttgacggaag aatgatgtgc 4500
ttttgccata gtatgctttg ttaaataaag attcttcgcc ttggtagcca tcttcagttc 4560
cagtgtttgc ttcaaatact aagtatttgt ggcctttatc ttctacgtag tgaggatctc 4620
tcagcgtatg gttgtcgctt gagctgtagt tgccttcate gatgaactgc tgtacatttt 4680
gatacgtttt tccgtcaccg tcaaagattg atttataatc ctctacaccg ttgatgttca 4740
aagagctgtc tgatgctgat acgttaactt gtgcagttgt cagtgtttgt ttgccgtaat 4800
gtttaccgga gaaatcagtg tagaataaac ggatttttcc gtcagatgta aatgtggctg 4860
aacctgacca ttcttgtgtt tggcttttta ggatagaatc atttgcacg aatttgtcgc 4920
tgtctttaaa gacgcggcca gcgtttttcc agctgtcaat agaagtttcg ccgacttttt 4980
gatagaacat gtaaategat gtgtcatcgc catttttagg atctccggct aatgcaaaga 5040
cgatgtggta gccgtgatag tttgcgacag tgccgtcagc gttttgtaat ggccagctgt 5100
cccaaacgtc caggcctttt gcagaagaga tatttttaat tgtggacgaa tcaaattcag 5160
aaacttgata tttttcattt ttttgcgtgt cagggatttg cagcatatca tggcgtgtaa 5220
tatgggaaat gccgtatgtt tccttatatg gcttttggtt cgtttcttcc gcaaacgctt 5280
gagttgcgcc tcctgccagc agtgcggtag taaagggtta tactgttgct tgttttgcaa 5340
actttttgat gttcatcgct catgtctcct tttttatgta ctgtgttagc ggtctgcttc 5400
ttccagccct cctgtttgaa gatggcaagt tagttacgca caataaaaaa agacctaaaa 5460
tatgtaaggg gtgacgcaa agtatacact ttgcccttta cacatttttag gtcttgcttg 5520
ctttatcagt aacaaacccg cgcgatttac ttttcgacct cattctatta gactctcggt 5580
tggattgcaa ctggtctatt ttctctttt gtttgataga aaatcataaa aggatttgca 5640
gactacgggc ctaaagaact aaaaaatcta tctgtttctt ttcattctct gtatttttta 5700
tagtttctgt tgcattggga taaagttgcc tttttaatca caattcagaa aatatcataa 5760
tatctcattt cactaaataa tagtgaacgg caggtatatg tgatgggtta aaaaggatcg 5820
gcggccgctc gatttaaaatc tcgagaggcc tgacgtcggg 5860

<210> 75
<211> 31
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 75
gagaggatcc ggaaggtgaa tcgaatttcg g 31

<210> 76
<211> 40
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 76
ctattgctgt cggcgctcat gattctccaa aaataatcgc 40

<210> 77
<211> 20
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 77
atgagcgccg acagcaatag 20

<210> 78
<211> 29
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 78
gaactctaga tcagaacgcc gccacggac 29

<210> 79
<211> 6591
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:Plasmid

<400> 79
gatccggaag gtgaatcgaa ttctggggct ttaaagcaaa aatgaacagc ttggtctata 60
gtggctaggt accctttttg ttttggacac atgtagggtg gccgaaacaa agtaatagga 120
caacaacgct cgaccgcgat tattttttgga gaatcatgag cgccgacagc aatagcaccg 180

acgccgatcc gaccgcgcat tggtcgttcg aaaccaaaca gatacacgct ggtcagcacc 240
ctgatccgac caccaacgcc cgggctctgc cgatctatgc gaccacgtcg tacaccttcg 300
acgacaccgc gcacgcgcgc gccctgttcg gactggaaat tccgggcaat atctacaccc 360
ggatcggcaa ccccaccacc gacgtcgtcg agcagcgcac cgccgcgctc gagggcggtg 420
tggccgcgct gttcctgtcg tcggggcagg ccgcggagac gttcgccatc ttgaacctgg 480
ccggcgcggg cgatcacatc gtgtccagcc cgcgcctgta cggcggcacc tacaacctgt 540
tccactattc gctggccaag ctcggcacgc aggtcagctt cgtcgacgat ccggacgate 600
tggacacctg gcaggcggcg gtacggccca acaccaaggc gttcttcgcc gagaccatct 660
ccaaccgcga gatcgacctg ctggacaccc cggcgggttc cgaggtcgcc catcgcaacg 720
gggtgccggt gatcgtcgcg aacaccatcg ccacgccata cctgatccaa ccgttggccc 780
agggcgccga catcgtcgtg cattcgccca ccaagtacct gggcgggcac ggtgcccga 840
tcgcgggtgt gatcgtcgcg ggcggcaact tcgattggac ccaggggcgc ttccccggct 900
tcaccacccc cgaccccagc taccacggcg tgggtgttcgc cgagctgggt ccaccggcgt 960
ttgcgctcaa agctcgagtg cagctgctcc gtgactacgg ctcggcggct tcgccgttca 1020
acgcgttctt ggtggcgagc ggtctggaaa cgctgagcct gcggatcgag cggcacgtcg 1080
ccaacgcgca gcgcgtcgcc gagttcctgg ccgcccgca cgacgtgctt tcggtcaact 1140
atgcggggct gccctcctcg ccctggcatg agcgggcca gaggctggcg cccaagggaa 1200
ccggggcgct gctgtccttc gagttggccg gcggcatcga ggccggcaag gcattcgtga 1260
acgcgttgaa gctgcacagc cacgtcgcca acatcgggtga cgtgcgctcg ctggtgatcc 1320
accggcatc gaccactcat gccagctga gcccgccga gcagctggcg accgggggtca 1380
gccccggcct ggtgcgtttg gctgtgggca tcgaaggtat cgacgatatc ctggccgacc 1440
tggagcttgg ctttgccgcg gcccgagat tcagcgccga cccgcagtcc gtggcgcgct 1500
tctgatctag acccgggatt taaatcgcta gcgggctgct aaaggaagcg gaacacgtag 1560
aaagccagtc cgcagaaacg gtgctgacct cggtatgaat tcagctactg ggctatctgg 1620
acaagggaaa acgcaagcgc aaagagaaaag caggtagctt gcagtgggct tacatggcga 1680
tagctagact gggcggtttt atggacagca agcgaaccgg aattgccagc tggggcgccc 1740
tctggtaagg ttgggaagcc ctgcaaagta aactggatgg ctttcttgcc gccaaaggatc 1800
tgatggcgca ggggatcaag atctgatcaa gagacaggat gaggatcggt tcgcatgatt 1860
gaacaagatg gattgcacgc aggttctccg gccgcttggg tggagaggct attcggctat 1920
gactgggcac aacagacaat cggctgctct gatgccgcgc tgttccggct gtcagcgcag 1980
gggcgcccgg ttctttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac 2040

gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac 2100
gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc 2160
ctgtcatctc accttgtctc tgccgagaaa gtatccatca tggtgatgc aatgcggcgg 2220
ctgcatacgc ttgatccggc tacctgccc ttcgaccacc aagcgaaaca tcgcatcgag 2280
cgagcacgta ctcggatgga agccggctctt gtcgatcagg atgatctgga cgaagagcat 2340
caggggctcg cgccagccga actgttcgcc aggtcaagg cgcgcacatgcc cgacggcgag 2400
gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc 2460
ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg 2520
ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg 2580
ctttacggta tcgccgtctc cgattcgcag cgcacgcct tctatcgct tcttgacgag 2640
ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgacgcc aacctgccat 2700
cacgagattt cgattccacc gccgccttct atgaaaggct gggcttcgga atcgttttcc 2760
gggacgccgg ctggatgate ctccagcgcg gggatctcat gctggagttc ttcgcccacg 2820
ctagcggcgc gccggccggc ccggtgtgaa ataccgcaca gatgcgtaag gagaaaatac 2880
cgcatcaggg gctcttcgc ttcctcgtc actgactcgc tgcgctcggc cgttcggctg 2940
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat 3000
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 3060
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacia aaatcgacgc 3120
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcggt tccccctgga 3180
agctccctcg tgcgctctcc tgttcggacc ctgccgctta ccggatacct gtccgccttt 3240
ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcgggtg 3300
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 3360
gccttatccg gtaactatcg tcttgagtc aaccgggtaa gacacgactt atcgccactg 3420
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcgggtg tacagagttc 3480
ttgaagtggg ggccctaacta cggctacact agaaggacag tatttggtat ctgcgctctg 3540
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 3600
gctggtagcg gtgggttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct 3660
caagaagatc ctttgatctt ttctacgggg tctgacgtc agtggaacga aaactcacgt 3720
taagggatth ttggtcatgag attatcaaaa aggatcttca cctagatcct tttaaaggcc 3780
ggccgcggcc gcgcaaagtc ccgcttcgtg aaaattttcg tgccgcgtga tttccgccca 3840
aaaactttta cgaacgttcg ttataatggg gtcatgacct tcacgacgaa gtactaaaat 3900
tggcccgaat catcagctat ggatctctct gatgtcgcgc tggagtccga cgcgctcgat 3960

gctgccgtcg atttaaaaac ggtgatcgga tttttccgag ctctcgatac gacggacgcg 4020
ccagcatcac gagactgggc cagtgccgcg agcgacctag aaactctcgt ggcggatcct 4080
gaggagctgg ctgacgagct gcgtgctcgg ccagcgccag gaggacgcac agtagtgagg 4140
gatgcaatca gttgcgccta ctgcggtggc ctgattcctc cccggcctga cccgcgagga 4200
cggcgcgcaa aatattgctc agatgcgtgt cgtgccgcag ccagccgcga gcgcgccaac 4260
aaacgccacg ccgaggagct ggaggcggct aggtcgcaaa tggcgctgga agtgcgtccc 4320
ccgagcgaaa ttttggccat ggtcgtcaca gagctggaag cggcagcgag aattatcgcg 4380
atcgtggcgg tgcccgcagg catgacaaac atcgtaaattg ccgcgtttcg tgtgccgtgg 4440
ccgccagga cgtgtcagcg ccgccaccac ctgcaccgaa tcggcagcag cgtcgcgcgt 4500
cgaaaaagcg cacaggcggc aagaagcgat aagctgcacg aatacctgaa aaatgttgaa 4560
cgccccgtga gcggttaactc acagggcgctc ggctaacccc cagtccaaac ctgggagaaa 4620
gcgctcaaaa atgactctag cggattcacg agacattgac acaccggcct ggaaattttc 4680
cgctgatctg ttcgacaccc atcccagact cgcgctgcga tcacgtggct ggacgagcga 4740
agaccgccgc gaattcctcg ctacactggg cagagaaaat ttccagggca gcaagaccgc 4800
cgacttcgcc agcgcttgga tcaaagaccc ggacacggag aaacacagcc gaagttatac 4860
cgagttgggt caaaatcgct tgcccgggtc cagtatgttg ctctgacgca cgcgcagcac 4920
gcagccgtgc ttgtcctgga cattgatgtg ccgagccacc aggcggcgcg gaaaatcgag 4980
cacgtaaacc ccgaggtcta cgcgattttg gagcgctggg cagcctgga aaaagcgcca 5040
gcttggtatcg gcgtgaatcc actgagcggg aaatgccagc tcatctggct cattgatccg 5100
gtgtatgccg cagcaggcat gagcagcccg aatatgcgcc tgctggctgc aacgaccgag 5160
gaaatgaccc gcgttttcgg cgctgaccag gctttttcac ataggctgag ccgtggccac 5220
tgcaactctc gacgatccca gccgtaccgc tggcatgccc agcacaatcg cgtggatcgc 5280
ctagctgata ttatggaggt tgctcgcattg atctcaggca cagaaaaacc taaaaaacgc 5340
tatgagcagg agttttctag cggacgggca cgtatcgaag cggcaagaaa agccactgcg 5400
gaagcaaaag cacttgccac gcttgaagca agcctgccga gcgccgctga agcgtctgga 5460
gagctgatcg acggcgctccg tgctcctctgg actgctccag ggcggtgccg ccgtgatgag 5520
acggcctttc gccacgcttt gactgtggga taccagttaa aagcggctgg tgagcgccca 5580
aaagacacca agggctcatcg agcctacgag cgtgcctaca ccgtcgctca ggcggtcgga 5640
ggaggccgtg agcctgatct gccgccggac tgtgaccgcc agacggattg gccgcgacgt 5700
gtgcgcggct acgtcgctaa aggccagcca gtcgtccctg ctctcagac agagacgcag 5760
agccagccga ggcgaaaagc tctggccact atgggaagac gtggcggtaa aaaggccgca 5820

gaacgctgga aagacccaaa cagtgagtag gcccgagcac agcgagaaaa actagctaag 5880
tccagtcaac gacaagctag gaaagctaaa ggaaatcgct tgaccattgc aggttggttt 5940
atgactgttg agggagagac tggctcgtgg ccgacaatca atgaagctat gtctgaattt 6000
agcgtgtcac gtcagaccgt gaatagagca cttaaggctc gcgggcattg aattccacg 6060
aggacgccga aagcttccca gttaaagtgc catctcgtag gcagaaaacg gttcccccg 6120
agggtctctc tcttggcctc ctttctaggt cgggctgatt gctcttgaag ctctctaggg 6180
gggctcacac cataggcaga taacgttccc caccggctcg cctcgtaage gcacaaggac 6240
tgctcccaaa gatcttcaaa gccactgccg cgactgcctt cgcgaagcct tgccccgcgg 6300
aaatttcctc caccgagttc gtgcacaccc ctatgccaaag cttctttcac cctaaattcg 6360
agagattgga ttcttaccgt ggaaattctt cgcaaaaatc gtccccctgat cgcccttgcg 6420
acgttggcgt cggtgccgct ggttgcgctt ggcttgaccg acttgatcag cggccgctcg 6480
atttaaattc cgagaggcct gacgtcgggc ccggtaccac gcgtcatatg actagttcgg 6540
acctagggat atcgtegaca tcgatgctct tctgcgttaa ttaacaattg g 6591